NORTH ATLANTIC TREATY
ORGANISATION

RESEARCH AND TECHNOLOGY
ORGANISATION

AC/323(IST-059)TP/262

www.rto.nato.int

RTO TECHNICAL REPORT **TR-IST-059**

A Framework for Network Visualisation

(Un cadre pour la visualisation
des réseaux)

Final Report of Task Group IST-059.

Published February 2010

Distribution and Availability on Back Cover

NORTH ATLANTIC TREATY
ORGANISATION

RESEARCH AND TECHNOLOGY
ORGANISATION

AC/323(IST-059)TP/262

www.rto.nato.int

RTO TECHNICAL REPORT **TR-IST-059**

A Framework for Network Visualisation

(Un cadre pour la visualisation
des réseaux)

Final Report of Task Group IST-059.

The Research and Technology Organisation (RTO) of NATO

RTO is the single focus in NATO for Defence Research and Technology activities. Its mission is to conduct and promote co-operative research and information exchange. The objective is to support the development and effective use of national defence research and technology and to meet the military needs of the Alliance, to maintain a technological lead, and to provide advice to NATO and national decision makers. The RTO performs its mission with the support of an extensive network of national experts. It also ensures effective co-ordination with other NATO bodies involved in R&T activities.

RTO reports both to the Military Committee of NATO and to the Conference of National Armament Directors. It comprises a Research and Technology Board (RTB) as the highest level of national representation and the Research and Technology Agency (RTA), a dedicated staff with its headquarters in Neuilly, near Paris, France. In order to facilitate contacts with the military users and other NATO activities, a small part of the RTA staff is located in NATO Headquarters in Brussels. The Brussels staff also co-ordinates RTO's co-operation with nations in Middle and Eastern Europe, to which RTO attaches particular importance especially as working together in the field of research is one of the more promising areas of co-operation.

The total spectrum of R&T activities is covered by the following 7 bodies:

- AVT Applied Vehicle Technology Panel
- HFM Human Factors and Medicine Panel
- IST Information Systems Technology Panel
- NMSG NATO Modelling and Simulation Group
- SAS System Analysis and Studies Panel
- SCI Systems Concepts and Integration Panel
- SET Sensors and Electronics Technology Panel

These bodies are made up of national representatives as well as generally recognised 'world class' scientists. They also provide a communication link to military users and other NATO bodies. RTO's scientific and technological work is carried out by Technical Teams, created for specific activities and with a specific duration. Such Technical Teams can organise workshops, symposia, field trials, lecture series and training courses. An important function of these Technical Teams is to ensure the continuity of the expert networks.

RTO builds upon earlier co-operation in defence research and technology as set-up under the Advisory Group for Aerospace Research and Development (AGARD) and the Defence Research Group (DRG). AGARD and the DRG share common roots in that they were both established at the initiative of Dr Theodore von Kármán, a leading aerospace scientist, who early on recognised the importance of scientific support for the Allied Armed Forces. RTO is capitalising on these common roots in order to provide the Alliance and the NATO nations with a strong scientific and technological basis that will guarantee a solid base for the future.

The content of this publication has been reproduced
directly from material supplied by RTO or the authors.

Published February 2010

ISBN 978-92-837-0084-5

Table of Contents

Chapter 4 – Hypernodes: Their Application to Visualisation of Networks 4-1

Annex D – Aspects of the Application of Information Theory to Visual Communication D-1

Annex E –Network Uncertainties E-1

Annex F – Survey Taxonomy and Analysis F-1

List of Figures

List of Tables

List of Acronyms

2-D	Two Dimensional
3-D	Three Dimensional
ACE	Alternating Conditional Expectation
ACM	Association for Computing Machinery
ACSC	Australian Conference on Computer Science
AGVis	Action Group on Visualisation
AG3	Information Visualisation Action Group
AGARD	Advisory Group for Aerospace Research and Development
ANB	Analyst's Notebook file format
API	Application Programming Interface
APVIS	Asia Pacific symposium on Information visualisation
AVI	Advanced Visual Interfaces
AZ	Arizona
BBN	Bayesian Belief Network
BELIV	BEyond time and errors: novel evaLuation methods for Information Visualization
C2	Command and Control
C2NetVis	Command and Control Network Visualisation
C3I	Command, Communication, Control and Intelligence
CA	California
CAN	Canada
CASOS	Computational Analysis of Social and Organizational Systems
CDC	Centers for Disease Control
CGI	Common Gateway Interface
CHI	Computer Human Interaction / Conference on Human Factors in Computing Systems
CIKM	Conference on Information and Knowledge Management
CMU	Carnegie Mellon University
CMV	Coordinated and Multiple Views
CNN	Cable News Network
COP	Common Operational Picture
COS	Chief Of Staff
COTS	Commercial Off-The-Shelf
CPU	Central Processing Unit
CSV	Comma-Separated Values
DA	Descriptive Aspects
DALO	Danish Defence Acquisition and Logistics Organization
DARPA	Defense Advanced Research Projects Agency
DC	Domain Context
DCIEM	Defense and Civil Institute of Environmental Medicine
DEU	Germany (Deutschland)
DIG-COLA	Directed Graph Layout through Constrained Energy Minimization
DIMACS	Center for Discrete Mathematics and Theoretical Computer Science
DNA	Dynamic Network Analysis
DNK	Denmark
DRDC	Defence Research and Development Canada
DRG	Defence Research Group
DSTO	Defence Science and Technology Organisation

ECU	Elementary Control Unit
EURO-DAC	European Design Automation Conference
FFI	Norwegian Defence Research Establishment
FGAN	Research Establishment for Applied Science (Forschungsgesellschaft für Angewandte Naturwissenschaften)
FKIE	The Research Institute for Communication, Information Processing and Ergonomics (Forschungsinstitut für Kommunikation, Informationsverarbeitung und Ergonomie)
FTE	Full Time Equivalent
FTP	File Transfer Protocol
GBR	United Kingdom
GEM	Generalized Expectation Maximization
GI	Graphics Interface
GML	Geography Markup Language
GPHIN	Global Public Health Information Network
GPU	Graphics Processing Unit
GRIP	Graph dRawing with Intelligent Placement
GUI	Graphical User Interface
GVF	Graph Visualization Framework
GXL	Graph eXchange Language
HAT	Human Factors, Applications and Technologies
HCI	Human Computer Interaction
HFM	Human Factors and Medicine
HQ	Headquarter
HTML	Hypertext Modeling Language
HTTP	Hypertext Transfer Protocol
I/O	Input/Output
I3C	Innovation, Initiative, Intuition and Creativity
IAEA	International Atomic Energy Agency
IBM	International Business Machines corporation
ICA	International Cartographic Conference
iCKN	innovative Collaborative Knowledge Networks
ICON	International Conference on Networks
IEEE	Institute of Electrical and Electronics Engineer
InfoVis	Information Visualisation
IP	Internet Protocol
ISCIT	International Symposium on Communications and Information Technologies
ISRI	Institute for Software Research International
IST	Information Systems Technology panel
ISTn	Information Systems Technology Technical Team “n”
IVC	Information Visualization CyberInfrastructure
J2	Intelligence and Security Affairs Sector
J6	Communications and Informatics Sector
J7	Education and Training Sector
J8	Resource Planning and Programming
JNDMS	Joint Network Defence and Management System
JUNG	Java Universal Network/Graph framework
KDD	Knowledge Discovery and Data Mining
KPI	Key Performance Indicator

Lab	Laboratory
LED	Light Emitting Diode
LNCS	Lecture Series in Computer Science
LS	Lambda Set
MILC	Multi-dimensional In-depth Long-term Case study
MIT	Massachusetts Institute of Technology
MST	Minimum Spanning Tree
MVC	Model-View-Controller
N/A	Not Applicable
N/X	Visualisation Network of Experts
NATO	North Atlantic Treaty Organisation
NC3A	NATO Consultation, Command and Control Agency
Net/NET	Network
NIH	National Institute of Health
NJ	New Jersey
NOR	Norway
NSF	National Science Foundation
NY	New York
OGDF	Open Graph Drawing Framework
Ops	Operations
PA	Pennsylvania
PARC	Palo Alto Research Center
PCT	Perceptual Control Theory
PDA	Personal Digital Assistant
PfP	Partnership for Peace
PoW	Programme of Work
R&D	Research and Development
RGB	Red, Green, Blue
RM-Vis	Reference Model for Visualisation
ROU	Romania
RSG	Research Study Group
RTA	Research and Technology Agency
RTG	RTO Task Group
RTO	Research and Technology Organisation
RWS	RTO Workshop
SIAM	Society for Industrial and Applied Mathematics
SIGCHI	Special Interest Group on Computer-Human Interaction
SNA	Social Network Analysis
SQL	Structured Query Language
SVG	Scalable Vector Graphics
SWE	Sweden
TAP	Technical Activity Proposal
TBD	To Be Determined
TCP	Transmission Control Protocol
TGRIP	Temporal Graph dRawing with Intelligent Placement
TM	Technical Memo
ToR	Terms of Reference

TR	Technical Report
TRADOC	Training and Doctrine Command
TTCP	The Technical Cooperation Program / Programme
TV	Television
UIST	User Interface Software and Technology
UK	United Kingdom
URL	Uniform Resource Locator
USA	United States
USD	United States Dollar
VA	Virginia
Vis	Visualisation
Vis N/X	Visualisation Network of Experts
VisTG	Visualisation Technology Group
VITA	Visual Interface for Text Analysis
VizDEC/DMSEC	Workshop on Visualization and Data Mining for Computer Security
VL	Visual Languages
VR	Virtual Reality
WG	Working Group
WHO	World Health Organisation
XGMML	eXtensible Graph Markup and Modeling Language
XML	eXtensible Markup Language

IST-059 Membership

CANADA

A. Bouchard (*)
Alain.Bouchard@drdc-rddc.gc.ca
+1 (418) 844-4000 x4744 (voice)
+1 (418) 844-4538 (fax)

DRDC Valcartier
2459 Pie-XI Blvd North
Val-Bélair, Quebec G3J 1X5
CANADA

Z. Jacobson
zack@sigmaxi.org
zack_jacobson@hc-sc.gc.ca
+1 (613) 946-9056 (voice)
+1 (613) 956-3166 (fax)

Health Policy Branch
Health Canada
Mailstop 1905C, Tunney's Pasture
Ottawa, Ontario K1A 0K9
CANADA

M.M. Taylor
mmt@mmtaylor.net
+1 (416) 488-2679 (voice)
+1 (416) 635 2191 (fax)

Martin Taylor Consulting
369 Castlefield Avenue
Toronto, Ontario M5N 1L4
CANADA

V.K.Taylor (Team Leader)
Vincent.Taylor@drdc-rddc.gc.ca
+1 (613) 993-9946 (voice)
+1 (613) 993-9940 (fax)

DRDC Ottawa
3701 Carling Avenue
Ottawa, Ontario K1A 0Z4
CANADA

J. Treurniet
Joanne.Treurniet@drdc-rddc.gc.ca
+1 (613) 990-7096 (voice)
+1 (613) 993-9940 (fax)

DRDC Ottawa
3701 Carling Avenue
Ottawa, Ontario K1A 0Z4
CANADA

DENMARK

L.M. Rasmussen (*)
lr@mil.dk
+45 72 57 15 52 (voice)
+45 72 57 51 20 (fax)

DALO
Lautrupbjerg 1-5
DK-2750 Ballerup
DENMARK
ATT FMT-C-TDBE12 Rasmussen, Lisbeth M.

GERMANY

A. Kaster (*)
a.kaster@fgan.de
+49 228 9435 492 (voice)
+49 228 9435 508 (fax)

FGAN-FKIE
Neuenahrer Str. 20
D-53343 Wachtberg-Werthhoven
GERMANY

M. Kleiber
kleiber@fgan.de
+49 228 9435 482 (voice)
+49 228 9435 508 (fax)

FGAN-FKIE
Neuenahrer Str. 20
D-53343 Wachtberg-Werthhoven
GERMANY

---

* Principal Delegate (National Leader).

NORWAY

J.T. Bjørke (*)
jtb@ffi.no
+47 3303 3837 (voice)
+47 3304 7834 (fax)

Norwegian Defence Research Establishment (FFI)
P.O. Box 115, NO-3191 Horten
NORWAY

ROMANIA

I. Bica (*)
ibica@mta.ro
+4021 335 4660 (voice)
+4021 335 5763 (fax)

Military Technical Academy
81-83 George Cosbuc Blvd.
75275 Bucharest
ROMANIA

UNITED KINGDOM

M. Varga (*)
mjvarga@QinetiQ.com
+44 (1684) 895712 (voice)
+44 (1684) 896502 (fax)

QinetiQ Plc
St. Andrews Road
Malvern, Worcestershire WR14 3PS
UK

UNITED STATES

M. Nixon (*)
Mark.R.Nixon@aero.org
+1 (310) 336-5783 (voice)
+1 (310) 336-6186 (fax)

The Aerospace Corporation
2350 E. El Segundo Blvd.
A3/2266
El Segundo, CA 90245
USA

A.K.C.S. Vanderbilt
drv@dramyvanderbilt.com
+1 (571) 723-5645

Vanderbilt Consulting
8101 Timber Valley Ct
Dunn Loring, VA 22027
USA

S.A.H. McMullen
sonya.mcmullen@techreachinc.com
+1 702-750-2936

8776 Country Pines Avenue
Las Vegas, NV 89129
USA

---

* Principal Delegate (National Leader).

A Framework for Network Visualisation

(RTO-TR-IST-059)

Executive Summary

IST-059 is the latest of a series of NATO research study groups investigating how to visualise effectively various kinds of datasets of interest in the defence community. In earlier groups, the question frequently arose of how to utilise visualisation to enhance network understanding, noting that proper application of visualisation tools and methods to the network domain could enhance the effectiveness of, for example, both military commanders and medical officers of health. IST-059 was formed to address this issue.

Given the large number of network presentation techniques available, both designers and users need assistance to discover the most effective technique for the particular task at hand. A framework was needed within which both the user's requirements and the properties of the available visualisation tools and methods could be matched. Frameworks for different aspects of the problem have been published, but all appear to be specific to a constrained domain. In an attempt to develop an overarching methodology, IST-059 designed its own framework, building on work of previous groups as well as work of the TTCP C3I Panel Action Group on Visualisation. The group also undertook a survey of tools and methods that addressed network visualisation issues, using a taxonomy which was developed for the purpose, with a goal to link the survey with the framework to facilitate selection of visualisation tools and methods that match a particular domain space and user role of interest.

The conceptual framework was used to "walk through" some test cases from different domains which provided insight into the most appropriate/effective linkage between the framework and the survey. What remains to be done, based on the results of the walk-through, is:

a) To refine the framework and the survey taxonomies to enhance the match between them;

b) To instantiate the framework;

c) To test the framework in realistic cases; and

d) To complete its integration with the survey.

As an unexpected outcome, some members of IST-059 believe the framework development, supported by underlying theoretical developments, is the start of a new "Unified Theory of Networks".

IST-059 operated through biannual business meetings and annual workshops, supplemented by continual electronic communication among its members. The workshops addressed identified visualisation topics of particular interest to the work of the group. The topics in successive workshops were: Social Network Analysis and Visualisation for Public Safety, Visualising Network Information, Network Analysis and Visualisation for Simulation and Prediction, and Visualising Network Dynamics. Each workshop provided an excellent forum for comprehensive interaction among group members and other international experts in an informal but structured setting. This interaction led to the establishment of some international collaboration that likely would not have occurred otherwise.

The short term cost avoidance due to collaborations directly attributable to the activities of the Group and its associated workshops has been estimated at some three million dollars; however the long term cost savings of these collaborations can not yet be estimated and may end up being much greater.

Un cadre pour la visualisation des réseaux

(RTO-TR-IST-059)

Synthèse

L'IST-059 est le dernier d'une série de groupes d'études de recherche de l'OTAN sur la façon de visualiser efficacement différentes sortes d'ensembles de données intéressant la communauté de Défense. L'IST-059 a été créé pour répondre à une question fréquemment soulevée au sein des groupes précédents : comment utiliser la visualisation pour améliorer la compréhension des réseaux (en notant qu'une application correcte des outils et méthodes de visualisation dans le domaine des réseaux pourrait améliorer l'efficacité, par exemple, des chefs militaires et des médecins de santé en même temps).

Compte tenu du grand nombre de techniques de présentation de réseaux disponibles, les concepteurs et les utilisateurs ont besoin d'assistance pour déterminer la technique la plus efficace pour cette tâche particulière. Il était nécessaire de déterminer un cadre où besoins de l'utilisateur et propriétés des outils et des méthodes de visualisation disponibles pouvaient concorder. Des cadres relatifs aux différents aspects du problème ont été publiés mais ils se sont tous révélés spécifiques à des domaines précis. Pour tenter de développer une méthodologie globale, l'IST-059 a conçu sa propre structure à partir des travaux des groupes antérieurs ainsi qu'à partir du travail du Panel TTCP C3I Groupe d'Action sur la Visualisation. Le groupe a également entrepris une étude des outils et des méthodes qui traitent des questions de visualisation des réseaux, en utilisant une taxinomie développée pour la circonstance. Le but était de relier l'étude au cadre afin de faciliter la sélection des outils et des méthodes de visualisation correspondant à un domaine particulier et à un pôle d'intérêt pour l'utilisateur.

Le cadre conceptuel a été utilisé pour « passer en revue » quelques cas de tests dans différents domaines fournissant un aperçu du lien le plus approprié/efficace entre le cadre et l'étude. Sur la base des résultats obtenus par la revue, il reste à :

a) Perfectionner le cadre et les taxinomies de l'étude pour améliorer la corrélation entre eux ;

b) Mettre à jour le cadre par des exemples concrets ;

c) Tester le cadre avec des cas réalistes ; et

d) Compléter cette intégration par une étude.

De plus, ce qui n'était pas prévu, quelques membres de l'IST-059 ont pensé que le développement du cadre appuyé par des développements théoriques sous-jacents, était le début d'une « Théorie Unitaire des Réseaux ».

L'IST-059 a fonctionné au travers de réunions d'affaires biannuelles et d'ateliers annuels, avec des contacts informatiques continuels entre ses membres. Les ateliers ont abordé des thèmes de visualisation choisis ayant un intérêt particulier pour le travail du groupe. Les thèmes abordés dans les ateliers successifs ont été les suivants : Analyse du Réseau Social et Visualisation pour la Sécurité Publique, Information sur la Visualisation du Réseau, Analyse du Réseau et Visualisation pour la Simulation et la Prévision et Dynamique sur la Visualisation du Réseau. Chaque atelier a organisé un forum de qualité, informel mais structuré, pour une interaction complète entre les membres du groupe et d'autre experts internationaux. Cette interaction a posé les bases de collaborations internationales qui n'auraient jamais existé autrement.

Les économies à court terme dues aux collaborations directement attribuables aux activités du Groupe et aux ateliers associés ont été estimées à environ trois millions de dollars. Les économies à long terme de ces collaborations ne peuvent pas être actuellement estimées mais finiront par être beaucoup plus importantes.

Chapter 1 – INTRODUCTION

1.1 OVERVIEW

Through visualisation, we attempt to understand our world. When we visualise a plan, a problem, a situation, or a structure, we are almost certainly visualising not only the objects and concepts, but also a network of relationships among them. Almost every action requires the actor to assess relationships. Even in such a trivial everyday action as to transport an object, the relationship between the object's size and the capacity of potential containers, the relationship between the terrain and the available transport mechanisms, the relationship between the cost of replacing the object and the risk of loss are only a few of the relationships that must be implicitly or explicitly considered.

Effective network visualisation tools and techniques can enable:

- A battlefield commander to visualise the relationships among his forces, between his forces and those of the enemy, between the forces and the landscape, between events at one time and events at another. All of these relationships constitute networks; understanding the networks through effective visualisation helps a commander gain information superiority.
- Designers and defenders of computer networks to visualise what is happening and what might be happening if certain events were to occur within those networks. Computers are valuable not just as processing centres, but as communicators that link people and ideas across the world. The network of their connections is part of our critical infrastructure and is vulnerable to attack, both physically and in the cybersphere.
- Intelligence personnel to assess potential terrorist threats by visualising the development and structure of networks of malicious groups; defenders can visualise the likely effects of specific interventions and non-intervention.
- Municipal officials to visualise the effects of natural and deliberate failures of part of the infrastructure on the behaviour of other parts, and to be able to see how the whole network is behaving at any time, both under normal conditions and in emergencies.
- Medical officers of health to protect against potential acts of bioterrorism, and to respond effectively to disease outbreaks.

The list of military and civilian requirements for network visualisation could be extended indefinitely, which hints at why, as of June 2008, over 500 different ways of representing networks had been collected on one Web-site [1], and why it may be hard to find the right technique to support effectively different user roles.

IST-059 is the latest in a series of research study groups that have been investigating the visualisation of massive military datasets of different kinds and from different viewpoints. In this work, the problem of how to present networks to aid the visualisation process frequently arose. IST-059 recognized that the large number of available network presentation techniques has provided a myriad of possibilities for designers. However, this plethora of options has given designers headaches in trying to determine which, among the many offerings, would be most effective to adopt or most sensible to adapt for the particular task at hand.

IST-059 observed that a missing key element was a framework within which both the user's needs and the properties of the available visualisation methods could be matched. Several frameworks for different aspects of the problem have been published, but none seemed to satisfy the overall requirement. IST-059 undertook to develop a general network visualisation framework – the Framework – that would aid the community in understanding their visualisation issues, thus leading to better design decisions. The group used the general visualisation “VisTG Reference Model”, originally published in [2], and the RM-Vis Framework for Visualisation developed by TTCP C3I AGVis [3] as foundation material. At the same time, the group undertook an extensive survey of tools and methods that address network visualisation issues – the Survey. The Survey was based on a taxonomy developed for the purpose which has the following main categories: Context, Network Representation, Analysis, Visual Enhancements, User Interaction, and Deployment. These and their sub-categories are described in Chapter 3. Although the Survey and Framework each stands on its own, IST-059 decided that it would be useful to be able to link the Survey with the Framework in such a way as to facilitate selection of visualisation technologies that are a good match for the problem space.

The IST-059 Framework structure is now largely designed and a worksheet has been produced to aid in analysing a problem in terms of its visualisation requirements. The part of the Framework that links the pattern of answers to a suitable display or display type is an on going activity; it is the intention of the next group to do this so as to link the answers to the technologies identified in the survey. The conceptual Framework has been used to “walk through” some test cases from widely differing domains and this work is giving insight to providing the right linkage between the Framework and the Survey. What remains to be done, based on the results of the walk-through, is:

a) To refine the Framework and the Survey taxonomies in order to provide a better match between them;

b) To implement the Framework so that it can be used effectively;

c) To test the Framework in realistic cases; and

d) To complete its integration with future versions of the Survey.

1.2 BACKGROUND TO THE RTG

Getting out from under the data flood and understanding its implications is a major problem for the military, as indeed it is for any up-tempo operation. DRG Panel 11, the Information Technology Panel under the old NATO Defence Research Group, considered that flexible and intuitive visual interfaces could contribute greatly to the effectiveness of interaction with the data flood and capability to extract and manage information from it. The Panel authorized a workshop titled “Visualising non-Visual Information” to help determine whether automated visual information processing was an appropriate domain for cooperative or collaborative investigation among the NATO partners.

An exploratory workshop, held in Brussels in November 1994, recommended that a Research Study Group be formed to help coordinate much needed research on data flood issues related to military and defence needs, and to maintain a broad overview and a detailed view of visual information management technologies. Such a Research Study Group (RSG) was created in 1996 with the aim of developing methods for presenting to human users the implications of the contents of large, complex and varying military-relevant datasets of diverse kinds. The RSG would operate by meeting semi-annually in the different member nations. Shortly thereafter, Panel 11 was decommissioned and the RSG was transferred to the Human Factors Panel, DRG Panel 8, as its RSG-30.

RSG-30 identified four modes of perception that support three classes of activity in which visualisation can play an effective role: monitoring, alerting, searching and exploring to support analysis, problem solving and briefing activities. In each of these areas, specific and different types of interaction with information are required, implying that different types of presentation may be needed.

RSG-30 agreed that "visualisation" implied understanding data in context as information, rather than simply displaying data on a screen. Other sensors besides the eyes are frequently valuable in creating a mental picture. Hence, although derived from an initial interest in discovering and displaying the content of massive textual datasets, the horizon of RSG-30 was expanded to include non-textual material. RSG-30 interpreted visualisation as a human activity, supported by technology, by which humans make sense of complex data. The Group considered visualisation support technologies, such as search engines, algorithmic processes and display devices and techniques, but only in relation to how they help humans to perform their tasks effectively. The Group emphasized the human use of the computational sub-system in ensuring that the right information be available in the form and at the time needed. In this, the Group's approach to the nature of visualisation is not inconsistent with that of the USA Army, which characterizes battlefield visualisation as "*the process whereby the commander develops a clear understanding of his current state with relation to the enemy and the environment, envisions a desired end state, and then subsequently visualises the sequence of activity that will move his force from its current state to the end state*" [4].

In 1996, RSG-30 formed an informal technical group known as the "Visualisation Network of Experts", or "Vis N/X", composed of known visualisation experts from the NATO countries. The group was to have an independent existence but be cognizant of, and responsive to, the needs of the RSG. Vis N/X was expected to conduct an annual workshop in conjunction with a regularly scheduled meeting of the RSG and was to be a sounding board and advisor to the RSG on developments in visualisation science and technology.

In 1998, while RSG-30 was still active, NATO's defence science and technology structure underwent reorganization, with the Defence Research Group (DRG) and the Advisory Group for Aerospace Research and Development (AGARD) being merged to form a new organization, the NATO Research and Technology Organisation (RTO). Following the reorganization, RSG-30 was retained as a RTO Task Group (RTG) under a newly formed Information Systems Technology (IST) Panel with the interim designation of IST-005, which later was changed to IST-013. The title of this RTG was "Visualisation of Massive Military Datasets".

The RTG confirmed the RSG-30 interpretation that visualisation was a human activity, supported by technology, by which humans make sense of complex data. With this understanding, the group developed the "IST-005 Visualisation Reference Model", which model, under its later name "the VisTG Reference Model" (Figure 1-1), has underpinned most of the group's subsequent work.

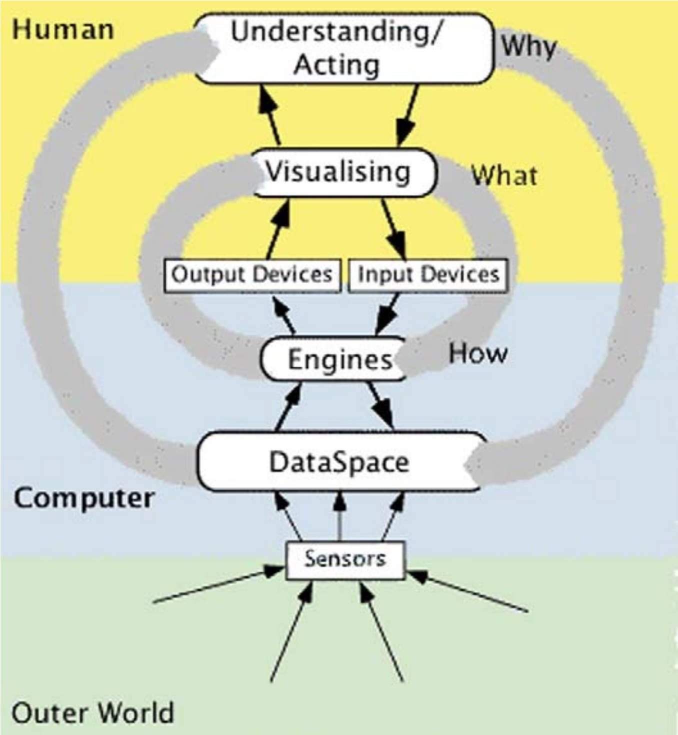


Figure 1-1: The VisTG Reference Model, Showing the Physical Devices that Enable Communication between the User and the Processes and Data in the Computer.

In June 2000 the Group delivered a workshop on "Multimedia Visualisation of Massive Military Datasets" in Quebec, Canada [5]. The workshop concluded: simple displays can be useful in dealing with complex data; modular and componentware structures ease system development; user involvement needs to continue throughout the development process; support for the user's innovation, initiative, intuition, and creativity (I3C) is important especially in the face of anomalous conditions; visualisation of relationships is an important unsolved problem; and evaluation is an important area of research. The importance of the interactive aspect rather than simply the presentation was brought out clearly throughout the workshop.

The RTG published its final report "Human Factors, Applications, and Technologies" – the HAT report [2] – in December 2000. This report, which included the "IST-005 Visualisation Reference Model", evaluated available technologies, applications for which visualisation technology might be useful, the probable value and difficulty of applying the technology to each application, and research requirements that promised to have the best payoff. The report was written to enable a potential user to see how existing or near-future technologies might apply to a real problem, or possibly to see that no existing technology provides a cost-effective solution. Likewise, it would enable a researcher to evaluate which research issues are key, having potentially high payoff in a number of areas, or which permit the direct possibility of implementing specific applications. The document also allows researchers and potential users in all the NATO countries to evaluate what and where work is being done, thus facilitating the development of synergistic efforts.

In 2001, IST-013 was succeeded by the RTO Task Group IST-021 "Multimedia Visualisation of Massive Military Datasets" [5]. The objective of this new Group was to evaluate and update the visualisation systems principles developed earlier and to deliver a workshop on Military Data Fusion and Visualisation [6], the results of which would help to guide the subsequent work of the group. Following this successful workshop, held in 2002 in Halden, Norway, the Group sponsored a workshop “Visualisation and the Common Operational Picture” in 2004 in Toronto, Canada [7].

In the wake of the events on September 11, 2001, IST-021’s parent body, the NATO Research and Technology Organization, tasked several of its technical groups to address problems of security and defence against terrorist attacks. Accordingly, IST-021 requested that the Vis N/X consider the subject of information visualisation needs for intelligence and counter-terror during its 2003 workshop, which it agreed to do. Presentations and discussions from this workshop can be found on the Vis N/X Web-site [8].

A major issue that came up repeatedly throughout the life of IST-021 was how to visualise networks effectively to aid in their understanding. The RTG recommended to its parent IST Panel that it should be succeeded by a RTG that would address this issue. The Panel accepted the proposal and the current RTG, IST-059 “Visualisation Technology for Network Analysis”, commenced work in April 2005 with a mandate to produce a report to further understanding of visualisation technology and techniques as applied to network analysis tasks. The report was to help identify where and how visualisation methodology might realistically benefit such tasks including using visualisation technology to discover relationships, present relationships and to analyse relationships within and across both structural and social networks. The Technical Activity Proposal (TAP) for IST-059 is included in Annex A.

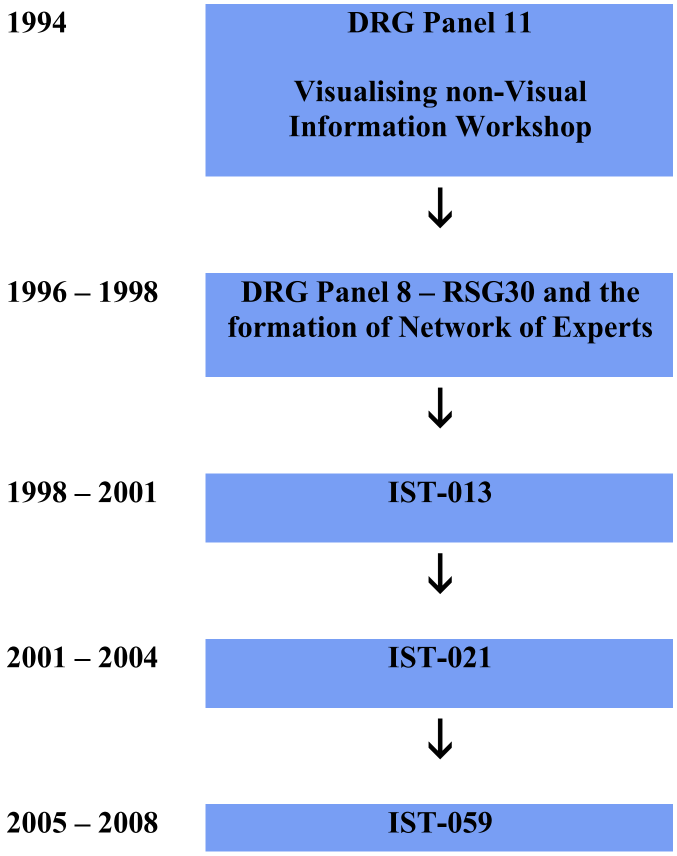


Figure 1-2: History of the Group.

1.3 VISUALISATION NETWORK OF EXPERTS (VIS N/X)

The idea for the Network of Experts came out of the 1994 NATO Defence Research Group (DRG) workshop "Visualising non-Visual Information", held at the Belgian Military Academy in Brussels. Some twenty five experts from several NATO countries met to discuss the state of the art in the emerging technologies supporting scientific, data, text and information visualisation, and make recommendations on how the DRG should support R&D. The workshop recommended that a Research Study Group (RSG) be formed to maintain awareness and to coordinate participating Nations' work in these fields with respect to military and defence needs. The workshop also recommended that the proposed RSG consider fostering an external independent network of visualisation experts that would be cognizant of, and responsive to, the needs of the RSG. This became the Vis N/X.

Vis N/X was initially created by the former NATO RSG-30 of DRG Panel 8 to be its informal technical advisory group and has continued to thrive under the patronage of succeeding visualisation Task Groups under the IST Panel of the RTO. The Vis N/X supports the RTG in its mandate. Since inception, the Vis N/X has held annual workshops in conjunction with meetings of the "parent" RTG except in years when the RTG itself delivered a formal NATO workshop.

The Vis N/X was a new concept in NATO research discussions and activities. In operation it offers an unofficial forum for researchers to exchange information, data and expertise. It carries some of the advantages that the NATO umbrella can offer, while avoiding some of the problems with more formal arrangements, including some Governments' occasional reluctance over the last decade to join in official arrangements.

Once the Vis N/X was constituted, its members decided that, in addition to normal e-mail interaction, they would endeavour to meet annually and would at the same time hold a workshop on a particular subject of interest to the associated Task Group. The timing and location of the annual meeting and workshop would be coordinated with the RSG and would be held in conjunction with one of its semi-annual business meetings. As much as possible, the Vis N/X would alternate its meetings between Europe and North America.

The Vis N/X has a select membership. It expands by inviting other experts identified by its members: any Vis N/X member may recommend an expert for membership provided that that expert comes from a NATO or PfP country. In June 2008, about 125 experts from 14 countries were members of Vis N/X. The Vis N/X expects to continue to operate through e-mail and to meet annually in conjunction with an RTG meeting to discuss topics of interest to the RTG; however it has not yet held its own separate meeting in a year in which the RTG is sponsoring an official NATO Workshop since its members are themselves likely to play key roles in such a workshop.

The Vis N/X has its own Web-site [8], which includes information on each of its workshops.

Table 1-1 summarizes the workshops held on behalf of the various visualisation RTGs, either under the NATO umbrella or by the Visualisation Network of Experts, Vis N/X.

Table 1-1: Vis N/X and *NATO* Workshops 1996 – 2008

Year and Location	Workshop Owner: Principal Topic
2008 **QinetiQ, Malvern, GBR**	Vis N/X: Visualising Network Dynamics
2007 **Aerospace Corp, El Segundo, USA**	Vis N/X: Analysis and Visualisation for Simulation and Prediction
2006 ***Danish Defence Research Establishment, Copenhagen, DNK***	***NATO IST-063: Visualising Network Information***
2005 **FGAN, Wachtberg-Werthhoven, DEU**	Vis N/X: Social Network Analysis and Visualisation for Public Safety
2004 ***Canadian Forces College, Toronto, CAN***	***NATO IST-043: Visualisation and the Common Operational Picture***
2003 **Penn State University, State College, USA**	Vis N/X: Information Visualisation Needs for Intelligence and Counter-Terror
2002 ***Army Logistics and Management College, Halden, NOR***	***NATO IST-036: Massive Military Data Fusion and Visualisation – Users Talk with Developers***
2001 **Åalborg University, Åalborg, DNK**	Vis N/X: Visualisation in Massive Military Datasets
2000 ***Defence Research Establishment Valcartier, Quebec, CAN***	***NATO IST-020: Visualisation of Massive Military Multimedia Datasets***
1999 **Defence Evaluation and Research Agency, Malvern, GBR**	Vis N/X: Information Visualisation
1998 **Defence and Civil Institute of Environmental Medicine, Toronto, CAN**	Vis N/X: Visualisation for Massive [Military] Datasets
1997 **Defence Evaluation and Research Agency, Malvern, GBR**	Vis N/X: Visualisation in Massive Military-Relevant Datasets
1996 **Consulting and Audit Canada, Ottawa, CAN**	Vis N/X: Visualisation in Massive Datasets

1.4 RTG PROGRAMME OF WORK

IST-059 considered various options for delivering its mandate and settled on three main work packages:

1) A survey of visualisation technology of potential relevance in network analysis: This task was to survey technology that is in production use as well as technology that is in the research and development stage within the individual countries. The RTG would attempt to identify in a broad

manner how the technology addresses the visualisation of one or more of: network structure; potential network behaviour; and actual and predicted network behaviour, particularly in the agreed application domains. The work that was carried out is described in Chapter 3.

2) Development of a network visualisation framework: This task would initiate development of descriptive and functional frameworks for network visualisation. A descriptive network visualisation framework should enhance an understanding of the commonalities of different ways of presenting network properties so that methods appropriate to one can be transferred to another. A functional network visualisation framework would characterize the interaction of a human operator with the network representation. This is discussed in detail in Chapter 2.

3) Develop and produce a Workshop on "Visualising Network Information": This task resulted in the delivery of a formal NATO workshop in October 2006 in Copenhagen, Denmark [9]. This workshop will be discussed further in Chapter 7.

In addition to the main work packages identified above, the RTG supported the Visualisation Network of Experts in delivering three workshops – "Social Network Analysis and Visualisation for Public Safety" in Bonn, Germany in 2005, "Network Analysis and Visualisation for Simulation and Prediction" in El Segundo, USA in 2007, and "Visualising Network Dynamics" in Malvern, England in 2008. The Vis N/X workshops are discussed further in Chapter 7.

As progress was made on the first two work items, it appeared that an integration of the survey and the framework might be feasible, which, if successful, could lead to the development of an automated visualisation analysis tool to support system developers/users. This is discussed further in Chapter 5.

The approved Terms of Reference and Programme of Work for the RTG are given in Annex A.

1.5 SUMMARY

IST-059 addresses the continuing requirement for the military and other up-tempo organizations to be able to visualise the implications of the ever-increasing amounts of data that modern technology makes available. Its predecessor RTGs had been concerned initially with the visualisation of textual data, then expanded with the problems imposed simply by massive amounts of data, and more recently with the evaluation of systems developed to aid personnel in visualising what their data/information means to their tasks.

Under the predecessor group IST-021 and its antecedents, a model for designing and evaluating visualisation systems was developed, called "the VisTG Reference Model." It was intended that this model provide the basis for tests in which different visualisation systems would be examined both inside the nation that developed them and in other nations. Several different systems were proposed by the nations, but for various reasons, very little was made available for testing.

The RTG and its predecessor groups, IST-013 and IST-021, sponsored NATO RTA workshops every other year in support of their missions – Valcartier, Canada in 2000; Halden, Norway in 2002; Toronto, Canada in 2004; and Copenhagen, Denmark in 2006. The proceedings of these workshops are available from the RTA [10][6][7][9].

To enhance the impact of the visualisation Task Group, in 1995 the group created a network of experts – Vis N/X which includes the members of the group plus invited experts. It is a voluntary organization, operating under the patronage and moral support of the extant RTG. So far, Vis N/X has held nine workshops, in both Europe and North America, the most recent having been held in November 2008 in the UK. These workshops were held annually until 2000 when the RTG started sponsoring a series of biennial NATO workshops.

Since then, the Vis N/X workshops have occurred in the out years, thus providing a series of annual workshops supporting the work of the RTG. Since 2002, the Vis N/X Workshops and the NATO workshops have followed a similar format, in which small working groups have addressed specific key questions, often suggested by the recommendations of working groups in earlier workshops.

1.6 REFERENCES

[1] Lima, M., (2008), "A Visual Exploration on Mapping Complex Networks" (online), http://www.visualcomplexity.com (Access Date: 15 June 2008).

[2] NATO IST-013/RTG-002, (2001), "Visualisation of Massive Military Datasets: Human Factors, Applications, and Technologies", *Technical Report, RTO-TR-030*, NATO Research and Technology Organisation.

[3] Bouchard, A. and Vernik, R., (2006), "Characterisation and Showcasing of Network Visualisation Approaches for Command and Control", In *Proceedings of Workshop IST-063/RWS-010 Visualising Network Data*, Copenhagen, p. 15, Copenhagen, Denmark, 17-20 October 2006.

[4] Department of the Army, (1995), "Battlefield Visualization Concept", *Technical Report, PAM 525-70*, Training and Doctrine Command (TRADOC), Fort Monroe, VA.

[5] NATO IST-021/RTG-007, (2007), "Multimedia Visualisation of Massive Military Datasets", *Technical Report, RTO-TR-IST-021*, NATO Research and Technology Organisation.

[6] NATO IST-036/RWS-005, (2004), "Massive Military Data Fusion and Visualisation: Users Talk with Developers", *Workshop Proceedings, RTO-MP-105*, Norwegian Defence Logistics and Management College, Halden, Norway: NATO Research and Technology Organisation, 10-13 September 2002.

[7] NATO IST-043/RWS-006, (2005), "Visualisation and the Common Operational Picture", *Workshop Proceedings, RTO-MP-IST-043*, Toronto, Ontario, Canada: NATO Research and Technology Organisation, 14-17 September 2004.

[8] Taylor, M., (2005), "IST-059 Network of Experts Home Page" (online), http://www.visn-x.net/ (Access Date: 4 July 2008).

[9] NATO IST-063/RWS-010, (2006), "Visualising Network Information", *Workshop Proceedings, RTO-MP-IST-063*, Copenhagen, Denmark, NATO Research and Technology Organisation, 17-20 October 2006.

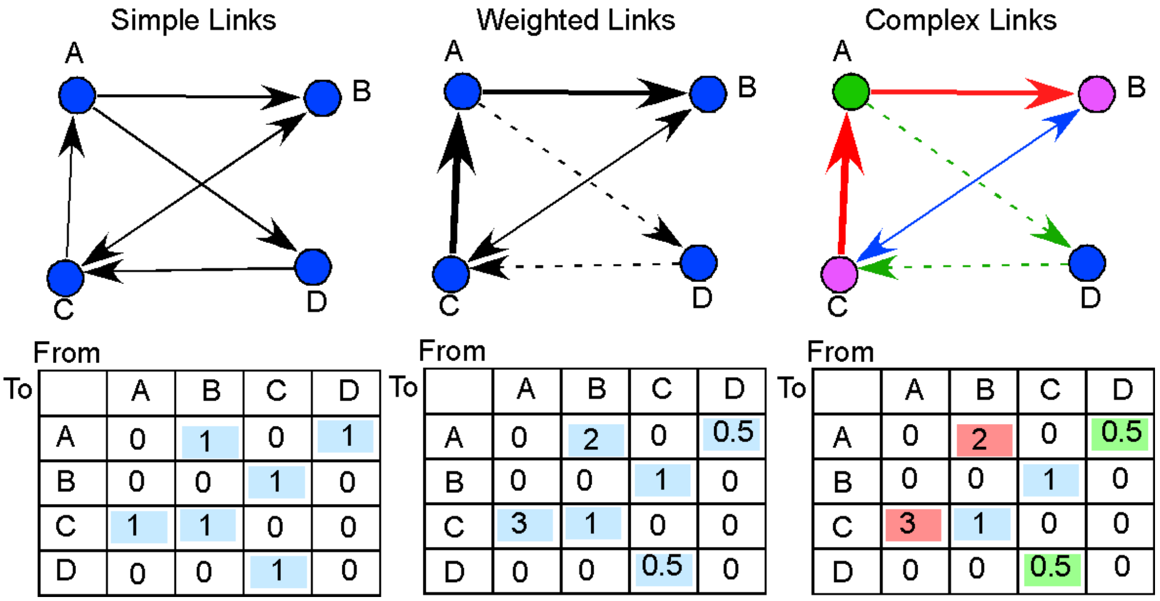
[10] NATO IST-021/RTG-007, (2002), "Multimedia Visualization of Massive Military Datasets", *Workshop Proceedings, RTO-MP-050*, Québec, QC, Canada: NATO Research and Technology Organisation. 6-9 June 2000.

Chapter 2 – A FRAMEWORK FOR NETWORK VISUALISATION

This chapter outlines the IST-059 Framework for Network Visualisation, and discusses why such a Framework is needed, what, in general terms, the Framework is, and how it may be developed into a generally useful form. More extensive technical detail and elaboration is in Annex B, for those who wish to delve more deeply into the issues.

2.1 WHY CREATE A FRAMEWORK FOR NETWORK VISUALISATION?

At first sight, there seems not much to say about visualising a network. Networks are often treated as though they were mathematical graphs. A graph can be specified as a matrix, in which occupied cells represent the links among the nodes that form the margins. In a pictorial display of a graph, nodes are often shown as little blobs, some of which are connected to others by lines to form a picture of all the connections. Perhaps the blobs or lines are coloured or sized differently to show properties of the nodes and links, but everything in the picture is also represented in the matrix, as in Figure 2-1. Looking at the picture, a person should be able to visualise how any one node connects to any other, see where there are nodes with many or with few links, and so forth.



From

To	A	B	C	D
A	0	1	0	1
B	0	0	1	0
C	1	1	0	0
D	0	0	1	0

From

To	A	B	C	D
A	0	2	0	0.5
B	0	0	1	0
C	3	1	0	0
D	0	0	0.5	0

From

To	A	B	C	D
A	0	2	0	0.5
B	0	0	1	0
C	3	1	0	0
D	0	0	0.5	0

Figure 2-1: A Network and its Representation (top) as a Node and Link Picture and (bottom) as a Matrix – (a, left) Simple Links; (b, middle) Links with a Weight Parameter; (c, right) Complex Links with Weight and Some Other Property Indicated by Colour.

If joining blobs with lines or filling cells of a matrix were all there were to the visualisation of networks, why would there be over 600 different kinds of network display (as of 2008/10/27) on a Web-site that showcases different kinds of network presentation (http://visualcomplexity.com)? Why did the IST-059 Survey of network visualisation applications (Chapter 3) list well over a hundred independent projects, even in its first quick scan of the field?

The answer is that the kinds of network that are interesting in real life consist of nodes that are much more complicated than simple blobs, and these complex nodes are connected by links that do much more than simply indicate whether or not one node is connected to another. The whole network is usually set in an environmental context that cannot be shown as a matrix. Different domains involve nodes of different types, and even with a single simple network, different tasks are made easier or more difficult by different kinds of presentation.

IST-059 recognized early that the very concept of a real world network was inadequately defined. Not only that, but although brilliant displays for many tasks involving networks had been devised, the members of IST-059 knew of no framework within which the tasks and the networks could be consistently described in a way that might help in designing displays useful for new problems. The need for such a Framework seemed evident.

IST-059 descended from earlier RTGs collectively known as VisTG. VisTG had considered the representation of massive military datasets, initially concentrating on the visual display of data that was inherently non-visual, such as the conceptual structure of text, which is, at heart, a network problem. Later, other kinds of dataset were considered. These had in common that large amounts of data needed to be displayed in a way that made sense to the user. Sometimes the data were dynamic, and the presentations had to change in real time. From the earliest days, however, it seemed that almost every problem involved at some point the display of relationships, in other words the display of networks.

Because of the wide variety of domains and tasks under consideration, IST-013 (one of the ancestors of IST-059) developed a three-pronged generic approach that could be used by someone interested either in selecting an existing display technology or in developing a display for a particular task [1]. The first approach was to create a functional model, the so-called “VisTG Reference Model” (see Annex H), which outlines the interactions that take place between a person wanting to visualise something and the computer in which the data is stored. The second was to develop a small taxonomy of data types and display types at a fairly abstract level, and to map data types to display types so as to suggest how effective displays might be designed. The third prong was to use a descriptive framework called “RM-Vis” ([2], and see Annex G), designed by a TTCP group, C3I AGVis (now C3I TP2).

The VisTG Reference Model (Figure 2-2) is conceptually based on the Perceptual Control Theory of W.T. Powers [3]. It consists of a three-level hierarchy of feedback loops: At the outer level, the user interacts with the dataspace, imagining and understanding its implications, and possibly influencing its content and structure. Psychologically, understanding comes by way of two complementary routes, logical analysis and visualisation. The VisTG Reference Model concentrates only on the visualisation route to understanding, explicitly ignoring logical analysis, while recognizing that displays must support analysis as well as visualisation.

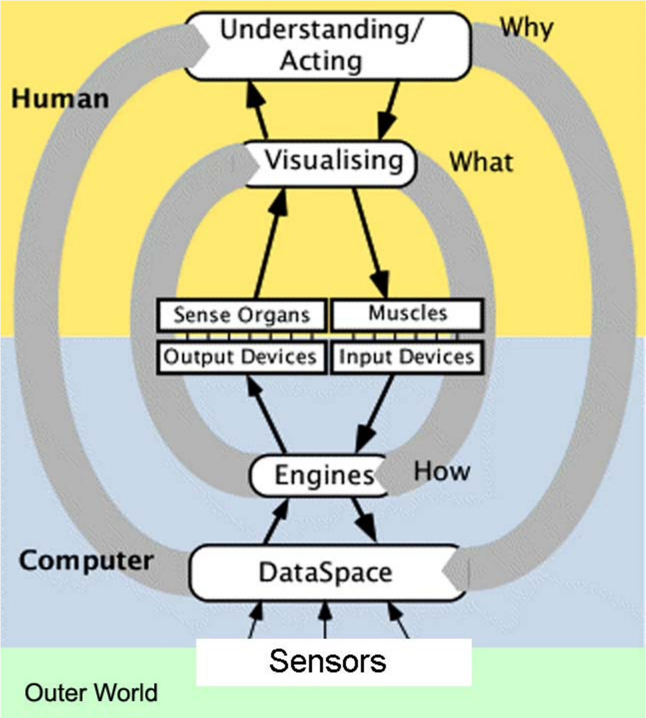


Figure 2-2: The VisTG Reference Model, Showing the Three Main Sets of Feedback Loops: Understanding–Dataspace, Visualising–Engines, and Physical I/O.

The user's understanding cannot interact directly with the dataspace, but must do so through some interpretive process, inside the human mind. Visualisation is one such process. It defines the middle loop of the VisTG Reference Model. The corresponding processes in the computer are known in the Model as "Engines". Engines select data from the dataspace, massage it, extract potentially useful structures and statistical properties, and organize the results both as modifications to the dataspace and for presentation to the user. In the language of the well-known Model-View-Controller approach to data presentation introduced with the SmallTalk computer language [4][5], the Model is in the dataspace, and different Engines produce the View and effect the Control. The VisTG Reference Model, though based on psychological theory, incorporates and extends the MVC structure.

The user's visualisation processes cannot interact directly with the Engines, but must do so through an innermost loop that implements two-way physical information transmission between human and computer using the Input-Output media, which usually include visual displays and keyboards and mice. However, output devices supporting visualisation may include anything perceptible by the user, including visible or audible language, haptic sensation, or other senses; input devices may be anything that the user can influence.

IST-013 used the VisTG Reference Model to generate a set of canonical questions, headed by "What state of the world do you want to be able to perceive" that should be addressed at each loop level when designing or choosing a display for a specific task.

The second approach taken by IST-013 was to develop a taxonomy of atomic data types and a corresponding taxonomy of display types. Users could identify what kinds of data were available that would permit them to create a display that would allow them to perceive whatever was the answer to the key question "*What do you* *want to be able to perceive*" of the VisTG Reference Model. Together, the VisTG Reference Model and the two taxonomies formed the basis of a functional Framework for visualisation of data in general [1].

Independently, a TTCP Group (C3I AGVis, now C3I TP2) created a descriptive framework for visualisation applications, which was given the name RM-Vis [2]. RM-Vis takes a quite different approach to that of the VisTG Reference Model, but one that complements the VisTG Reference Model very nicely. A description of RM-Vis is in Annex G. It defined four domains, loosely depicted as orthogonal axes in a space of description. Using this descriptive space, the TTCP group created a database of visualisation applications. IST-059 used the structure of this database as a basis for its own Survey of Network Visualisation projects and applications (Chapter 3).

The RM-Vis framework considered what had to be displayed and for whom, and included a descriptive axis for types of display, but did not consider the activities that go on when a user is trying to understand a dataspace through visualisation. IST-059/RTG-025 believed that combining the two approaches might be fruitful in developing a useful Framework for Network Visualisation.

2.2 DEFINING A FRAMEWORK

Different authors have different ideas as to what constitutes a Framework for visualisation. For example, Schulz and Schumann [6] treat the stages in processing from database to interactively produced display as a Framework. Within this Framework are different types and styles of display suited to different user requirements. Their entire Framework, however, is contained in just one of the four stages depicted in Figure 2-4 (below). The VisTG Reference Model is also a process framework, but one that incorporates user interaction and some cognitive functioning as well as display output. Others, such as RM-Vis, have taken descriptive taxonomies, of tasks, of graphical representations, of data, to constitute Frameworks. Chapter 3 notes yet other views, with perhaps a stronger emphasis on the computational side. There is no consensus, so we must define the term for the purposes of the work of IST-059.

For the work of IST-059, a Framework would assist a user to determine a display type that would suit the task at hand using the available data, and to discover whether there is an available application that would produce that kind of display from the data. It inherently involves the development of taxonomies, and at the same time it treats the stages of processing. Since the domain of interest is the display of networks, a large part of the work of IST-059 concerned the description of networks and the tasks for which different kinds of user might need to display some aspect of a network. Here we give a sketch of that work, which is described in more depth in Annex B.

2.2.1 Networks are More than Just Graphs

Networks are sometimes considered to be the same as graphs, sets of points connected by lines. They are not. Graphs are mathematical objects, whereas networks exist in the real world.

Networks are of two kinds, physical and conceptual. Physical networks are embodied in tangible structures, such as the set of wired interconnections of computers on an Ethernet, or the connections of power sources and sinks that constitute the electric supply infrastructure of a country or continent. Conceptual networks may have a physical substrate, as discussed below under "Embedding Fields" ("Embedding Field" is a concept introduced in this report; see Section 1.2.4 below, and Annex B, Section 1.2.1), or they be independent of any physical substrate, as, for example, the interrelations of the factors that influence a commander's plan of action, the likenesses among documents and intercepts that alert an intelligence officer to a potential opportunity or danger, the syntactic and semantic relationships that define two different networks over the words of a text, or the social connections that underlie the spread of ideas or of diseases.

Neither kind of network, physical or non-physical, is just a graph. A graph is a mathematical abstraction of a real-world network, eliminating messy real-world considerations, and ignoring any task-relevant context or embedding fields. In a graph, a link simply connects two nodes. In a network, two nodes may be connected by links of different types, nodes can have different roles, and the meanings of nodes and links may depend on the environmental and task context.

2.2.2 Global and Local Attributes of Networks

Networks come in different flavours, and have many properties not captured by graphs. Some kinds of network are:

- **Point-to-Point** – The classic network represented by a graph or matrix (e.g. Figure 2-1). Nodes are defined and each node is or is not linked to each other node by a link with some "weight". Nodes and links may have internal structure and processing capabilities. Different kinds of node may play different roles, and different kinds of link may connect nodes in a variety of ways. These aspects are not captured by graphs.
- **Broadcast** – A Broadcast Network must support traffic between transmitting and receiving nodes, usually but not necessarily over a continuous medium. A transmitting node cannot know which, if any, of many eligible receiving nodes may receive the traffic (e.g. airborne infection). Broadcasting is often a property of a sub-net of a network that is mostly point-to-point. There are two types of Broadcast Network, Ephemeral and Stigmergic.
  - In an "**Ephemeral Broadcast**" network, traffic not received at the time of transmission is lost. The adjective "Ephemeral" is usually omitted, and a "Broadcast Network" is taken to be ephemeral unless the context suggests otherwise.
  - In a "**Stigmergic Broadcast**" network, "traffic" is left in the environment and may be received at an indeterminate later time by an indeterminate number of receivers (e.g. ruts that tend to guide later traffic through a muddy field, or the clues left by a criminal that are read by a detective). A Stigmergic Broadcast network is often simply called a Stigmergic network.
- **Fuzzy** – Entities are not well defined as being nodes or links. Nodes may be somewhat linked to other nodes (e.g. suitability of road for heavy traffic). The membership of an entity in the class "node" or "link" may depend on the user's purpose. "Fuzzy" should not be confused with "probabilistic" or "stochastic".
- **Striped** or **Multimodal (Coloured)** – In a striped network, Nodes of type A can be linked only to nodes of type B (e.g. humans and malaria-carrying mosquitoes). Striped networks are a special class of multimodal network. In a multimodal or coloured network a node belongs to one of a range of classes or roles.

Links can have weights or strengths, but several different properties equally might deserve to be called the "strength" or "weight" of a link, some of them simultaneously:

- **Utilization** – If the link is of a kind that has traffic, how much traffic does it carry?
- **Capacity** – How much traffic could the link sustain?
- **Availability** – What is the probability the link will be open for traffic?

- **Coherence** (of a traffic-free link) – How tight is the relationship between the connected nodes? (*sibling is tighter than second cousin; "see" is more closely related to "view" than to "grow"*)
- **Fuzzy Membership** – How much like a link is the connection?

Links may be directed or undirected, elementary or bundled (compound). For example, person A might at the same time:

- be **the father of** person B;
- **lend** money to B;
- **enjoy** B's company; or
- **telephone** B **frequently**.

These attributes have obvious implications for display. A bundled link is a candidate for drilling down to examine its elementary constituent links, whereas an elementary link is not. The display should be capable of indicating to the user which is the case.

Nodes also can have a variety of properties, but these relate largely to their functions of transforming patterns of one or more inputs into one or more outputs, with varying delays. In this, nodes are like software functions, and there is no obvious set of properties with which to categorize them, other than by the dynamics of whether outputs are emitted synchronously with inputs, after a fixed or random delay, probabilistically or definitively, by whether the output is a sustained or impulsive effect of an input, and by whether the node emits output of the same type as the input.

Network displays often compress sub-nets and show a whole sub-net as a simple node. Any node that non-trivially transforms its inputs into its outputs is such a compacted sub-net; it is like a software routine that can be displayed as a block in a flow diagram. The display should probably allow the user to determine whether a node is compacted and is therefore likely to contain information that might be of use.

Nodes within a network also may differ in some of the properties assigned above to the network as a whole. For example, some nodes may broadcast, while others are point-to-point; some traffic may be ephemeral while some is stigmergic. These factors do affect the requirements for display in support of the user's visualisation, but not in a way that has yet been incorporated into the Framework.

Many mathematical properties can be computed from a graph, particularly from the graph abstraction of a network. Several such properties are developed in Annex C. Frequently used in the analysis of social networks, among many others, are:

- **Network Topology**: e.g. random, scale-free, tree.
- **Centrality**: Distribution of linkage degree over the nodes, distribution over the nodes of the likelihood a path between other nodes passes through a particular node, and so forth.
- **Directivity**: Whether links are unidirectional or two-way.
- **Cyclicity**: Is there a path over links from one node through other nodes and back to the original?
- **Diameter:** The longest geodesic between any pair of nodes (a geodesic is the shortest path between two specified nodes).

The mathematical properties of fuzzy networks are less well developed than those of crisp networks, but should reduce to those of crisp networks in the limit of binary membership functions (in which only zero or unity membership values are allowed).

2.2.3 Fuzziness and Uncertainty

Fuzziness and probabilistic data (or uncertainty) are often confused. However, the difference between them is easily illustrated: If I know that John is 188 cm tall (6 ft. 2 in.), and I am asked "Is John tall", I may answer "pretty tall". His tallness is a fuzzy property. On the other hand, I may be told that Bill is "rather short", which to me means that his height is probably somewhere around 165 cm, and unlikely to be more than 175 or less than 155. His height is a precise number that I know only as a probability distribution. Bill's height is not fuzzy, though his membership in the category "short" is fuzzy. It is Bill that is fuzzily "short", Bill's height that is probabilistically known to me.

In a network, nodes and links can have probabilistic properties. A road that passes over a lift bridge is not always a link between the two sides of the bridge, but at any moment it is clearly a link or not a link. On a wider time scale, it is a link with a probability less than unity of being available. On the other hand, a road that is subject to traffic jams may be clearly a link when traffic is moving freely, less of a link in dense traffic, and be hardly be a link at all when traffic is moving at a crawl. It has a fuzzy membership in the class "link", with a membership value that varies from near unity at times when the road is clear to near zero when there is a static traffic jam.

Real-world entities can have fuzzy membership not only in the class "link" but also in the class "node", as illustrated in Figure 2-3.

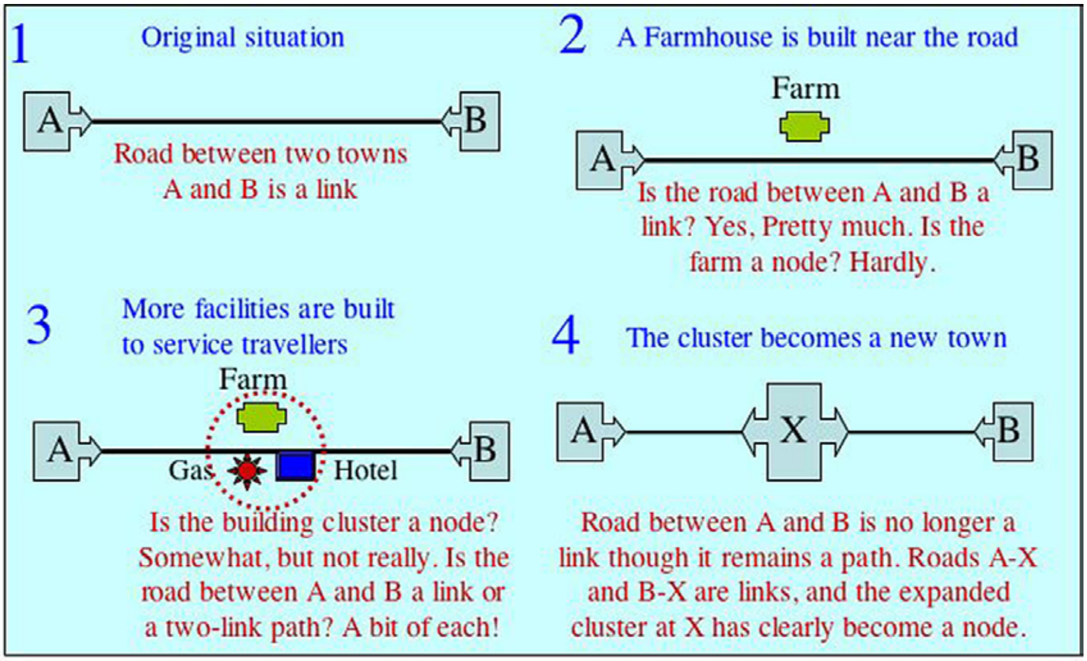


Figure 2-3: The Link Between A and B Becomes a Two-Link Path as the Membership of the Building Group Increases its Fuzzy Membership in the Class "Node".

The figure shows four stages in the historical development of a change in the structure of a network. Initially, there are two towns, A and B, connected by a road that is a link between the two nodes A and B. Over time, the road stays the same, but its status as a link between A and B changes, as a cluster of buildings takes shape around a farmhouse. The cluster of buildings becomes more and more like a node (increases its fuzzy membership in the class "node"), and the road becomes more and more like a two-link path A to X to B, as the membership in class "link" of the A-B connection decreases and the membership in class "link" of the A-X and X-B connections increase.

Uncertainty is quite different. In the fuzzy example above, the observer has no uncertainty as to the nature of the road or of the cluster of buildings that eventually developed into town X. The road is what it always was, and the properties of the buildings are perfectly known. The fuzziness becomes important when considering the mathematical and analytic properties of the network. When the road is somewhat a link between A and B and somewhat a two-link path, an analysis based on the crisp existence or absence of a link (as opposed to a path) will fail.

Uncertainty that matters is uncertainty of the user about some aspect of the network relevant to the task at hand, especially if that uncertainty affects a decision about some action. The user's uncertainty has two main sources: imprecision in the acquisition or analysis of the data, and the user's inability to be sure of an interpretation of precise data. In different circumstances, either kind of uncertainty may dominate. Often both work together. The user's ability to interpret precise data may be diminished either by the display of too little data or by the cluttered display of too much data. Presentation of the "right" amount of data is an issue that has been addressed using information theory (Annex B, Section 1.3.2), especially for the display of maps of road networks (Annex D).

Display presentation may affect the user's ability to interpret, but it can never influence imprecision in the data to be displayed. Presentation technique may, however, affect the user's ability to factor the imprecision of the data into the interpretation of the displayed data. The problem is that to show data imprecision both takes up display real-estate and is likely to divert the user's attention from the representation of the data themselves.

What attributes of uncertainty might be usefully displayed? The Network of Experts (N/X) Uncertainty working group at the 2007 El Segundo meeting listed: Reliability, Confidence, Accuracy, Precision, and Consistency. These attributes refer to different components in the train that leads to confidence in a decision.

- Reliability refers to the source of data and the route between that source and the data as displayed. It has a historical background, since a source cannot be known to be reliable or unreliable from one report. Only after several reports have been received and their data checked against other data from the same or different sources can the reliability be assessed.
- Confidence may refer to the confidence of the source reported along with the data or to the confidence of the user in the data or in the implications of the data.
- Accuracy might refer to the correctness of the data as compared to the real-world truth, but since this can never be ascertained, it is not a very useful construct. It is possible, however, to assess the likely range of deviation of a particular datum from what might be the result of other measures of the same thing, and it is not unusual for a measure to be given as x ± y.
- Precision refers to the likelihood that successive measures of the same thing result in similar data. Both Accuracy and Precision are more readily considered in connection with an attribute that has a scalar or vector value than in connection with the structural attributes of a network.

- Consistency refers to the repeatability of a datum based on different observations of the same thing. In a network context, this might include such things as whether A and B are likely to be connected by a link on Tuesday if they were so connected on Monday. In this sense, Consistency has a wider range of application than does Precision. Consistency may also refer to how well the implications of one set of observations agree with the implications of another set.

Other than Confidence, all these suggested attributes of uncertainty relate to the provision of the data for presentation to the user. Those attributes might, in principle, be computed and displayed to a passive user. To display confidence, however, requires that the user let the computer know the appropriate level of confidence about something displayed, or about its implications. This is possible only with Interactive, Coordinated, or Mediated context of use (see Section 2.3.2 for definitions of these terms).

The concept of “uncertainty’ goes hand in hand with that of “information”. Information is the reduction of uncertainty in the user’s mind. Information theory has been used to design and evaluate displays, from the viewpoint of minimizing the user’s uncertainty about the task-relevant aspects of the data. Some of these information-theoretic approaches to display design and evaluation are discussed in the following section, as well as in [7], and Chapter 4 and Annexes B, D, and E of the present report.

2.2.3.1 Uncertainty, Entropy and Information

Entropy and information are related but distinct concepts. Shannon [8] defined “information” as the reduction of uncertainty in the receiver about some aspect of the transmitter of a message. The receiver starts with some probability distribution over the set of messages the transmitter might send, and after a message has been received, has a different probability distribution over the set of messages, this new probability distribution referring to the receiver’s uncertainty as to which message was actually sent. The difference between the initial uncertainty about which message might be sent and the final uncertainty about which message was actually sent is the information transmitted.

Shannon’s uncertainty is formally the same as entropy, though used in conceptually different ways. Both are based on the same simple formula summing over $p \log p$, where p is the probability or probability density of a possible state. Entropy can be computed from any suitable collection of probabilities, and could be inherent in the structure of a network. Uncertainty, on the other hand, is normally “about” something. Uncertainty implies communication or observation, and in this context “messages” are observations of the transmitter. Since the transmitter is simply the object of observation, the concept of information transmission applies without modification to observation of the world, or of a display.

Shannon’s area of application was telephone communication, and the central construct of his work was the idea of communication channel capacity: how much information cold can be communicated through a noisy connection from a transmitter to a receiver. In the case of observation of a real-world network through a channel that passes through sensor systems to a computer database, from the database through algorithms to some abstraction, from the abstraction to a display, from the display to the user’s visualisation, and from the visualisation to the user’s understanding of the real-world situation, there are many opportunities for information loss and many different entropies to consider. Overall, the user’s task usually requires only a very small amount of information from the real world, such as “would breaking this link cause significant damage”, “which person is the leader of that group”, or “is that object a nuclear facility”? The entire reason for creating the long, heterogeneous channel is to transmit that small quantity of information from the real world to the user’s understanding.

Although the information to be transmitted is usually very limited, many of the stages in the transmission channel are of high entropy, most of which is irrelevant to the user's purpose, as suggested in Figure 2-4. The world observed by the sensor systems and entered in the dataspace is of very high entropy, but only a portion of it is displayed. With luck, all of the task relevant information is still available in the display (Figure 2-4 suggests that it may be not quite what is in the world, nor in the dataspace). From the display and from prior knowledge and skill (shown as "Memory" in Figure 2-4), the user develops a moderately high entropy visualisation of the part of the real world that contains the task-relevant data, and from the visualisation and analysis, the user extracts the low-entropy task relevant information.

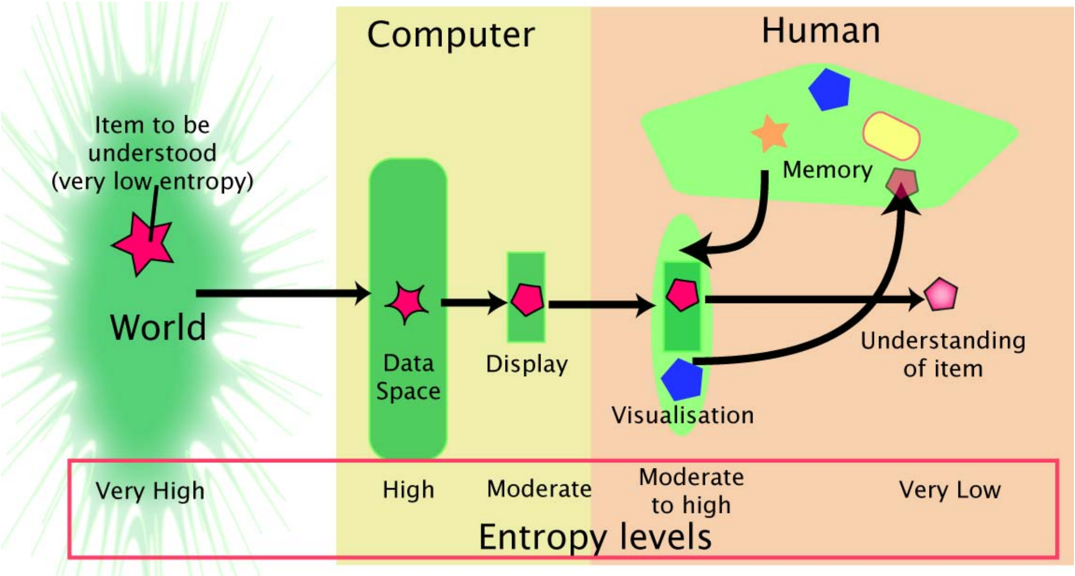


Figure 2-4: Schematic Showing Changes of Entropy as the User Obtains a Small Amount of Task-Relevant Information from the Real World, by Way of Sensor Transfer to the Dataspace, Selection and Algorithmic Manipulation to Form a Display, Visualisation Augmented by the User's Prior Knowledge, and Finally Understanding Based on Visualisation.

Since the sensor systems cannot know what aspect of the world interests the user at a given moment, the dataspace becomes filled with much that is irrelevant. Both the world and its abstraction in the dataspace contain information about what interests the user, though the representation in the dataspace may not contain everything about the item in the world, and what is represented may be distorted, as suggested in Figure 2-4. The same applies to the transformation between the dataspace and the display, though if the display is well designed, a higher proportion of its structure is devoted to the item of interest to the user. The user's visualisation may well be of higher entropy than the display, because the user contributes background knowledge, and that knowledge might well be able to fill in aspects of the item of interest that were lost in the preceding information channels. The final result of all this high-entropy high-information processing is the construction of a very low entropy representation of the item of interest in the user's mind.

2.2.4 Embedding Fields

Although a graph can exist *sui generis*, a network exists only in some real-world context. That context gives meaning to the network above and beyond its mathematical properties. To display something of the context usually helps a user to understand the implications of a display, but at no time can all the context be displayed

– it would be the entire universe! The concept of an “embedding field” helps to define the context likely to be useful. The “embedding field” is an important theme of the IST-059 Framework, and is a concept that we have not seen described elsewhere.

The “embedding field” of a network, from the viewpoint of the user and for the purposes of the task at hand, is the context within which the network functions. There are two different kinds of embedding field: a support structure within or over which the network is defined (a “semantic embedding field”), and an environment within which the network functions, but which does not itself support the network (a “pragmatic embedding field”).

The concepts of “semantic” and “pragmatic” embedding fields are analogous to the way those terms are used in linguistics. In linguistics, “semantic” relations are among the words in a text, according to their type and normal usage, whereas “pragmatic” relations are to states and events outside the text. For example, “Theodore Roosevelt voted for Julius Caesar as President of Nigeria in 1532” is semantically acceptable, but pragmatic nonsense.

IST-059 considers embedding fields to be important because they are the context for, and may provide meaning to, operations on the network itself. Frequently they constrain the possible behaviour of the network, and often they are the reason why a user wants to visualise the network. Accordingly, the display of a network often will include display of some embedding field. When the appropriate embedding field is effectively included along with the display of the network itself, the meaning of the network to the user can be greatly enhanced.

The concept of an “embedding field” was initially suggested by a pair of hypothesized assertions:

1) A physical network always has the possibility that a conceptual network lies on top of it. The conceptual network may map homologously onto the physical network if the relationships between nodes are defined as such, but in most cases, the conceptual network involves only sub-sets of the physical network.
2) A conceptual network may exist without any underlying physical network.

Examining these assertions led to the concept of the embedding field for a network, regardless of whether it has a physical substrate.

A network in the real world consists of physical or conceptual entities connected by relationships that may be

- Physically embodied (e.g. roads, wires); or
- Purely conceptual (family tree, social influence, conceptual relationship, etc.).

A network may be *embedded* in a physical or conceptual substrate, but what determines its “embedding field” for the purposes of display is the set of contextual attributes in which changes make a difference to the network ***from the viewpoint of the user and for the user’s current purpose.*** An active embedding field can be thought of as the currently relevant context. It may be semantic or pragmatic.

For example, a road network exists in a landscape of hills, valleys, rivers, towns, viewpoints, places of archaeological interest, and so forth. Exactly where between towns a road is laid may make no difference to a traveller, but it does make a difference to the people who live and work near the roads. For the traveller uninterested in the view, the embedding field may consist simply of the choice points and travel distances, elements of a semantic embedding field; for the local inhabitant or the tourist photographer, it is likely to include the geographical landscape, a pragmatic embedding field. The pragmatic embedding field is environmental, and not supporting.

A supporting (semantic) embedding field for a network is often another network on which it lies:

> *e.g. for a contagious disease, the network of infections is embedded in the network of social contacts, but this is not true for an airborne disease or one with an insect vector.*

Networks can inherit properties from supporting embedding fields:

> *e.g. location from a geographic embedding field, potential infectivity contacts from a social contact network embedding field.*

A supporting type of embedding field constrains the properties of the embedded network, but new attributes can be developed:

> *e.g. contacts are limited to those of the embedding social network, but contact type – casual, intimate, telephonic, etc. – may be attributes of the network of interest.*

A point-to-point network can have a broadcast network as an embedding field:

> *e.g. battlefield radios can be heard by friend and enemy, but the message traffic may be encoded so that a message can be received only by the intended recipients.*

Non-physical networks may have a physical substrate, and vice-versa. The network of links among pages on the World-Wide-Web (the Web) is non-physical, but it exists on a substrate network of information packet transmissions. The packet transmission network is itself non-physical, being defined by the possibilities for message traffic among computers running the appropriate protocol software; this protocol-based network depends on a physical network that consists of computers interconnected by physical wires or radio links.

Take another example. The network of possible airline connections derivable from published schedules is non-physical, but to implement a trip using the scheduled connections requires a network in which the nodes are airports and the links are defined by the traffic of physical aircraft. The aircraft travel according to a published schedule, but with variations due to events not forecast in the schedule plan. There is also a non-physical network in which the links between the airport nodes are defined by the actual trips taken by all passengers on a given day. The trip-based network has as its substrate a physical network whose links are defined by the actual aircraft flights between airports. This physical network itself depends on a non-physical conceptual network defined by the schedule plan. The difference between the trips planned on any given day and those actually taken (as affected by delays and cancellations) indicates the importance of the intervening physical network. Here we have a case of a non-physical network that has a physical network substrate, which in turn is based on a non-physical network.

As the examples show, embedding fields may be hierarchically nested. In the network of Web pages, one level of embedding field has its links defined by the two-way passage of *http* protocol messages between the servers and the clients. This network depends on a traffic-free conceptual network defined by the targets of the links specified on Web pages. At a lower level, there is a network of TCP and IP protocol connections among machines whose software is configured appropriately. Below that there is a network of computers that can communicate by fixed link or wirelessly.

All these embeddings serve to support the network of Web-page traffic. But not only do they support the Web network, the same protocol-based network also supports a completely independent network of e-mail message connections, and the e-mail network in its turn supports, in part, the social network of relationships among people – a network that is separately supported by other communication networks, such as the telephone system and the physical transport system.

Embedding fields, whether semantic or pragmatic, are hard to avoid, but are seldom noticed in discussions of network visualisation, which so often are limited to the display of graphs. Embedding fields are the context in which a network exists and that gives the network its meaning to the user.

2.2.4.1 Embedding Fields of the Display Medium

Not only the network, but also the display medium can be considered as a hierarchy of embedding fields, the root of which is the set of pixels of the display screen, intermediate levels might be 2-D and then 3-D spaces containing objects, while the leaves might consist of the coloured lines and objects used to show the network attributes of concern.

It is reasonable to speculate that the immediately ancestral embedding field for the display of the network may be the appropriate environment in which to display the user-relevant contextual embedding field of the network.

2.2.5 Dynamic Aspects of Networks

A network can change over time on all scales from its global structure to the movement of units of traffic over a single link. Consider, for example, the passenger transportation network. Two hundred years ago it had no railways and what little intercity travel there was went by road; one hundred years ago most intercity travel was by a vast rail network, and there was no air traffic; starting perhaps fifty years ago, rail lines began to be torn up, and more and more travel was by road and air; possibly in the future, the trend will be reversed and rail will again take over from increasingly expensive road and air travel.

In this example case, the network structure changes in a way that can be described as a smooth transition if one considers only the global parameters. On a finer scale, however, the changes might sometimes have been more abrupt. One day a town is served by a train and the next the service is gone. Travel shifts to the road or to the air if that is an option.

The above example suggests two kinds of change that may affect a network, changes in the structural linkages as rail lines are built and removed, and changes in the traffic patterns within a fixed structure, as would happen over time between rush hour and the dead of night. Other kinds of change may also be important for different user tasks. Changes in global attributes, such as density, variance of centrality, modularity, can be important. For example, when a hierarchic terrorist organization reconfigures itself as a distributed cellular structure, the change can be manifest in several global attributes, and it may well be more important for the authorities to visualise the implications of those global changes than to see the details of the organization.

Changes of this kind can profoundly alter the behaviour of a network. For example, altering the link density of social contacts might mean the difference between an infection fading out and the infection becoming a global pandemic. The same applies if the “infection” is an idea, a meme, which might be stifled at birth or might grow to become a worldwide religion, depending on the density and strength of the social links through which it is transmitted.

Either of the cases mentioned above can occur in an acyclic network. Oscillations, feedback reinforcement, and chaotic behaviour cannot. Accordingly, if a small change in link structure introduces cycles into a previously acyclic network, the dynamic behaviour of the network might change dramatically. A person may hear of a new idea from one friend and pass it on to another without much thought, but if she hears it again from a second source, the idea might well seem more plausible and important. The second source, however, might easily be repeating what had originally come from that same person by way of the first friend, completing a positive feedback loop.

If the link density is high, this kind of reinforcement can easily lead to the development of independent sub-net modules in some of which the idea is taken as truth, and in others it is taken as clearly false. Much political and religious conflict is likely to be due to such behaviour in cyclic networks. The question is how to present to a user the critical aspects of these global parameter variations, as well as the actual or predictable concrete consequences of the changes.

At the smallest scale, the important change might be that a particular packet of data was or was not transmitted between two significant nodes. If such a situation occurred, it is hardly likely that any generic display would lead the user to see it. The display either would have to be focused on the specific link, which could be problematic if there were many such links in a given network, or would have to be an alerting display that would call attention to the event (or non-event).

When a network is treated as a graph, there is no embedding field, but sometimes changes in the relationship between a network and one of its embedding fields might be important. For example, consider the network defined by the links among pages on different Web-sites. When a client follows a link, packets are transmitted over an embedding network of physical links between computers. If one of the computers that is used by a high proportion of the packet traffic between a particular client and server has been compromised, the compromised data does not affect the Web network, but it does affect the users of that client-server link. Removal of that computer from the embedding network would not change the Web network, though it might influence the responsiveness of some requests. What it would do is make the users less liable to real-world effects of illicit use of the intercepted data.

For the user, interesting changes may be retrospective, ongoing, or prospective. It may be important for the user to see that change did happen, resulting in a new stable state (Explore mode perception), that change is ongoing (Monitoring or perhaps controlling by actively influencing the ongoing change), or how change is likely to evolve (again Explore mode perception). The kinds of display suited to these three possibilities are likely to be quite different, just as they are for the various spatial scales of change discussed above.

2.2.6 What is a Framework: Summary

A Framework for Network visualisation should include typologies or taxonomies, not only of networks, but also of tasks, of display techniques, of user roles, and so forth. Many such are discussed in Chapter 3 and Annex B. But a Framework should include more than a list of taxonomies; it should include a procedure for using those taxonomies to allow a user to select an application appropriate for the task at hand, or to allow a developer or researcher to see the need for some new development – a new application, or perhaps a novel method of display. A Framework not only guides a user to what is available. It points the way to what ought to be made available.

2.3 THE IST-059 FRAMEWORK FOR NETWORK VISUALISATION

The IST-059 Framework for Network Visualisation builds on both the VisTG Reference Model and the RM-Vis descriptive framework. From the VisTG Reference Model it derives the procedures that help a user to select displays suited to the task at hand, and from the RM-Vis Framework it derives domains of description both for the data and for the user and task.

However, to take advantage of the Framework most effectively, the user should be able to find applications that are capable of reducing the data at hand to the display forms that are likely to be appropriate to the task. For this, IST-059 initiated a Survey of available software libraries and applications for network display (Chapter 3). The Survey elements are based on the RM-Vis structure, which should allow software to be developed that would allow the Framework procedure to serve as a front-end or interface through which the user would interact with the Survey. Even without the Survey, however, the Framework should assist the user in determining what kinds of displays might be useful for the task at hand, and should assist the developer and researcher to identify task types for which adequate displays should be made available.

Figure 2-5 suggests the place of the Framework in the workflow of visualising a network. Figure 2-5a shows the normal workflow, which starts with a task that concerns a real-world network. Somehow, information about that network has been captured in an abstract form in a computer database. Some of those data are used by algorithms to produce computed properties of all or part of the network Those properties are then manipulated by display technologies and some aspects of them made available to the user, who uses them to visualise something about the real-world network that is the real point of the task.

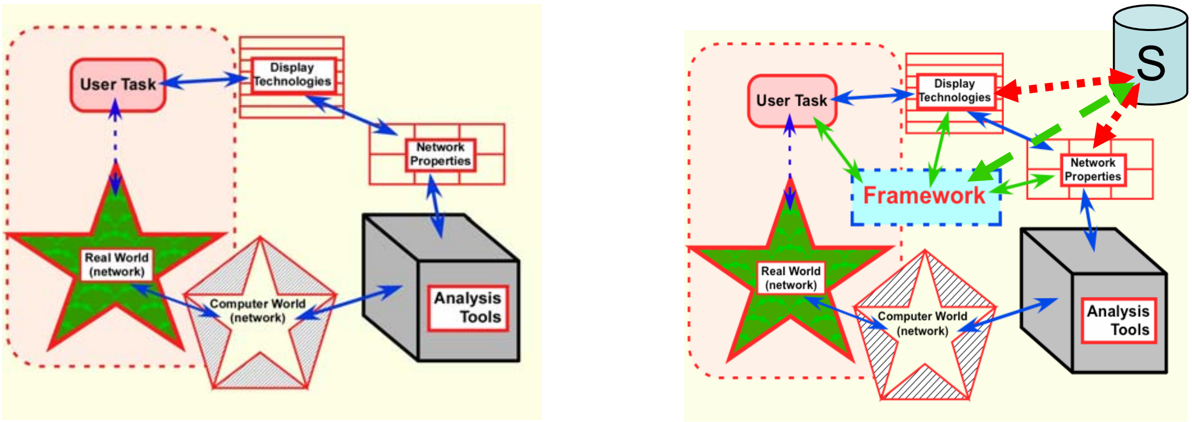


Figure 2-5: The Framework and the Workflow in Visualising Network Data – (a, left) the normal workflow; (b, right) the elements with which the Framework interacts.

Figure 2-5b suggests how the Framework and the Survey relate to this flow. The flow of abstractions does not change, but the Framework and Survey assist the user to choose and to manipulate its components. The user defines the task in some appropriate way. The Framework then is used to assess what kinds of display technologies used with what network properties will most readily allow the user to visualise what the task demands. The Framework can then be used to query the Survey to determine whether applications exist that will accept those properties and create the appropriate displays.

As currently envisaged, the Framework would not address the selection of analysis tools used to compute the desired network properties, but it is not unreasonable to suggest that it might be extended to do so.

The Framework is not simply a structure that relates the elements of the computational flow, as suggested in Figure 2-5. It is also a process that a user can follow. That process is based on the VisTG Reference Model, and is rooted in the question “What state are you trying to achieve”, followed by “How does that state differ from the existing state” and “What do you need to see in order to answer the question?”

When talking about visualisation, these questions usually resolve into “What do you want to be able to visualise for your task that you cannot at the moment visualise?” In the context of network display, this can be translated as “What properties of the network would you want to see that you can not now see, in order to help with the real task?” Both these questions imply a presumption that users have some definable knowledge of how their current understanding differs from the understanding they would like to achieve.

2.3.1 Modes of Perception

As described in the Final Report of IST-013 [1] following Taylor [9][10], four modes of perception can be categorized according to when and why the perception is used:

- *Monitoring and controlling* modes use perceptions of changing states of the world for real time purposes, either passively to maintain situation awareness or to ensure that the observed states remain within tolerable limits, or actively to influence them to approach desired conditions.
- *Searching* also serves real time purposes. It supports monitoring and controlling when data are lacking in the monitored state, by looking actively for the missing information. The data are used when found.
- *Exploring* is a background activity that does not support real-time monitoring and controlling. Information is acquired about states and structures of the world that are unlikely to change very much by the future time when the information may be useful for real-time monitoring and controlling. When the need eventually arises, prior exploration will have obviated the need for at least some real-time search.
- *Alerting* differs from the other three in that it is a highly parallel background process, and in humans likely to be non-conscious and automatic. In computer systems, alerting is likely to be supported by daemons that monitor the dataspace. The user specifies conditions or states that might suggest a requirement or an opportunity for monitoring or controlling, or that may signal the possible termination of a Search. Humans have evolved comparable internal autonomous alerting systems. An everyday example from human vision is the rapid eye-flick that often follows an unexpected motion in the visual periphery. The eye-flick allows the person to assess whether the movement signifies something that should be watched, without much distracting from whatever was in focus at the time. Likewise, one readily hears one’s own name in a conversational hubbub, while other names go unheard. In computerized systems, alerts can be set so that when an automated process detects a specified pattern in the data, an output triggers one of the human alerting systems. For example, when one of the daemons has detected the existence of the condition it was set up to notice, a portion of the visual display might blink or be shown in an unusual colour, or the sound pattern of an ongoing process might change.

These four modes presuppose different kinds of answer to the basic question of the VisTG Reference model “What properties of the network would you want to see that you can not now see, in order to help with the real task?” and are likely to suggest quite different displays of the same data.

Monitoring or Controlling: If the user is monitoring or acting to influence some developing situation in real time, the interesting properties are likely to be the ones that are changing and that are relevant to the primary task. The bandwidth of the display, and particularly the user’s ability to interpret it, are of prime importance. Bjørke and Varga (Chapter 4) and Bjørke (Annex D) use an information-theoretic approach to this question in the context of automating the presentation of varying-scale maps of networks such as roads. Their approach should be generalizable to other situations, and seems to be closely related to the information-theoretic approach to display provided by Smestad [7] in the Final report of IST-021. In general, if the user is monitoring or controlling, the presumption is that much of the context is already in the user's head from previous use of Explore mode perception (see below), and the display of real-time variation need only show so much of it as to make clear where the changes fit into, and how they alter, the relatively static context.

Searching: If the user is searching, the nature of what is sought is likely to be known. It may be possible to define the missing information in terms suited for automation, and thus to set up an alerting daemon to aid in the search by marking plausible regions of the dataspace wherein the desired information might be found. If not, then the display probably has to support not only the search, but also the ongoing monitoring process on behalf of which the search is being performed. The issue is likely to become one of linking in the user's mind the monitoring display and the search display with which the user navigates the dataspace. In terms of the VisTG Reference Model, the search is a supporting loop in the hierarchy, in which the user may actively control the Engines of the display system.

Exploring: When the user is Searching, some particular piece of information is being sought; when he or she is Exploring, the idea is to develop an understanding of the general context in which future monitoring or controlling may take place. In particular, effective exploration at one time should reduce the need for search at a later time. In a mundane example, Search takes the question "I need a pencil and must find one" to guide actions that eventually result in seeing a pencil in a drawer. Exploration takes the question "What is in those drawers" or "Where are there pencils" and notes that a pencil is in one particular drawer, so that later the "I need a pencil" state is immediately resolved by the knowledge of where the pencil is likely to be found. In the context of display, Search is likely to concern localized properties of the dataspace, whereas Exploration is likely to concern more general structural properties.

Alerting: By its very nature, Alerting is a background activity carried on by autonomous processes not involved in whatever the user is doing at the moment. It does not involve the construction of any displays, except when an alerting event is detected. Then, the display the user is currently using must be modified to indicate the existence of the alert, presumably by using one of the human's internal alerting systems to draw attention to the situation. Ideally, this modification allows the user to devote minimum attentional resources to determining whether the alert truly signifies something of immediate interest, and if it does not, should allow the user to return quickly and easily to whatever was in progress before the alert occurred.

The IST-059 Framework process asks the user which mode of perception is foremost. Does the user want to visualise the changing state of something for controlling or monitoring – in a historical review, monitoring is possible though controlling is not – or is the matter of interest to determine the structure of a network so that, say, its possible future behaviours may be understood more readily when they occur (Exploring)? Is the user looking for a particular aspect of the network, such as a key node or a region of potential vulnerability (Searching)? Perhaps the user only wants to see where in a network structure certain conditions are met as a prelude to further activity, or to be notified when incoming data match certain criteria (Alerting). The display requirements are rather different in all these cases.

Table 2-1 suggests influences of the perceptual mode on the display and the user's action.

Table 2-1: Perceptual Modes and Probable Display and Interaction Consequences

Perceptual Mode	Appropriate Display	Interaction
Monitoring/Controlling	Focus on network attribute being monitored or controlled, with context in background.	Monitoring: Possibly navigation. Controlling: Navigation and any methods of influencing dataspace.
Searching	More even display, perhaps with some increased detail near centre of area being searched. Focus on components of attribute being sought.	Navigation only. Includes informational zoom and navigation in attribute space, not just geographical space.
Exploring	Same as Searching, but perhaps with less concentration on specific attributes.	Same as Searching.
Alerting	No display until alerting condition found. Then minimally intrusive alerting indicator associated with area currently in user's focus.	Ability to shift easily to new focus on situation that led to the alerts, and more importantly, ability to revert, dismissing the alert if false alarm or unimportant.

2.3.1.1 Informational Aspects of the Modes of Perception

Because of the different uses of the modes of perception, they impose different informational requirements on the channels from real world to user understanding (Figure 2-4).

When the user is Monitoring or Controlling, attention is focused on the dynamic variations of one or two aspects of the data in the dataspace. These may be varying because they reflect a varying real world, because they are in a display of a dynamic simulation, or because the user is manipulating the dataspace interactively. However that may be, the requirement is that the user be continuously supplied with information about the current state of the attributes of interest. The information rate depends on the required precision, but the data source is defined and delimited. The display therefore can be of low instantaneous entropy, but with a bandwidth determined by the rate of change of the attributes being monitored.

In strong contrast to Monitoring or Controlling, Exploring is done to develop as much background knowledge as possible about relatively stable aspects of the dataspace (or the real world). The display may be complex and of high bandwidth, to allow the user to Explore without unnecessary attention-taking navigation through the dataspace. But it need not be continuously available. Indeed, by its nature, Exploring is a background activity, always susceptible to interruption on behalf of more pressing Monitoring or Controlling that could be signalled by an Alert. Exploring therefore supposes a high-entropy display, but with an arbitrarily low average bandwidth.

Searching shares in some attributes of Exploring and some of Monitoring/Controlling. Search is always done in support of an ongoing Monitoring/Controlling activity that requires some currently unavailable data. Search, therefore, requires a channel that is available when required, in contrast to Exploring, which can use a channel whenever time is free. If the result of a Search is delayed, the effective bandwidth of the Monitoring/

Controlling channel is automatically reduced. Search, on the other hand, benefits from reducing the need for the user to navigate through the dataspace, and therefore from a more complex display than is normally useful for Monitoring/Controlling. Search therefore seems to require a relatively high-entropy display with a high average bandwidth.

Finally, Alerting normally requires no display at all. Alerting processes (daemons) operate entirely within the computer, and require their own access to the dataspace, but that access is in parallel with the processes relating to user interaction with the dataspace. Only when an alerting event occurs does an Alerting daemon require access to the display, and then it requires only enough to allow the user to direct a Monitoring/ Controlling activity to the portion of the dataspace that triggered the Alert. However, that access usually needs to be immediate. Alerts therefore require a low average bandwidth channel with high availability but low bandwidth (a simple flashing light might sometimes be sufficient, for example).

These requirements are summarized in Table 2-2.

Table 2-2: Informational implications of Modes or Perception

	Availability	Instantaneous Entropy (Display Complexity)	Average Bandwidth
Monitoring/Controlling	High	Low	Low
Searching	High	High	High
Exploring	Low	High	Low
Alerting	High	Very Low	Very Low

2.3.2 Context of Use

A network may be displayed for one user or for several simultaneously. It may be manipulated in real time or be viewed as a static picture. It may be used by an end-user for discovery or to brief an audience on matters of interest. All of these possibilities have implications for what is displayed and how the display is controlled. In addition to the four perceptual modes, we recognize four different viewing regimes: Interactive, Coordinated, Mediated, and Passive. In the first three modes, the display is altered while it is being used.

- In Interactive mode, a single end-user manipulates the display and the database in real time. This is the canonical situation reflected in the VisTG Reference Model of Figure 2-2.
- In Coordinated mode, more than one end-user observes the display, and more than one user has responsibility for altering its content. The coordination among the users is an issue, and only one of them can be controlling any one aspect of the display at a given moment.
- In Mediated mode, one person, whom we may call an operator or a presenter, interacts with a display on behalf of the end-user. A lone user such as a commander might ask the operator to change the display in this way or that; in a briefing situation any of the viewers may be able to ask the person doing the briefing about aspects of the displays. In either case, there is interaction between the mediator and the user(s), as well as between the mediator and the display.
- In Passive mode, the user observes the display without influencing it in real time. An unlimited number of users can observe any particular display in passive mode. The display itself may change under the influence of an operator, but the users have no influence on the operator.

Table 2-3 suggests which perceptual modes are most likely to be used under different circumstances. Some modes are not applicable under some circumstances. A single user cannot be working coordinated, as coordination implies that more than one user is actively observing and influencing the display; multiple users viewing simultaneously cannot all be interactively controlling the display or its content; and if multiple viewers look at the display at different times and places, the display is very probably static, which implies passive viewing. Most often, the only effective perceptual mode for passive viewing is Explore. The user looks to see what can be discovered about what is displayed, and expects it to remain valid for some time thereafter.

Table 2-3: Perceptual Modes Most Likely to be Used in Different Circumstances

	Interactive	Coordinated	Mediated	Passive
Single End-User	All Modes	N/A	Explore, Search	Explore
Multiple Users Viewing Simultaneously	N/A	Monitor, Explore, Search, Alert	Explore	Explore
Multiple Users Viewing Separately	N/A	Monitor, Explore, Search, Alert	Explore, Search	Explore

If there are multiple viewers, either they are co-located, viewing the same display, or they are separately viewing individual displays. In the Passive case, the viewer has no influence on the display. Though the different users may have different views on the same dataspace, what happens is no more under the viewer's control than is the TV program in a standard broadcast. Much the same is true in Mediated viewing, since control of the display is vested in a single operator. However, if the viewers are not co-located, each may be able to control a separate View on a common Model (using the Model-View-Controller language). Control of the View is what allows multiple users viewing separately to use the Search mode of perception. Finally, in the Coordinated case, each may be controlling a separate View on a common Model, and more than one may be controlling the common Model that all are Viewing. Interaction among the multiple viewers then becomes a critical aspect of the display system design, if not of what is presented on the different screens.

2.3.3 The Worksheet

The Framework process begins by inducing the user to specify the task requirements, quite possibly thereby helping the user to clarify what actually is wanted to satisfy those requirements. In many cases, the user has at first only a vague idea of what actually is wanted, and finds it hard to translate this vague idea into a set of criteria for the selection of a suitable application or display technique. To this end, a worksheet was drafted, setting out a series of questions for the user to answer. A draft worksheet is shown in Chapter 5 in the form of a spreadsheet with some example problem statements filled in, though in fully implemented form it would probably be a Web-based form that eventually serves as a query interface to the Survey, allowing the user to explore the problem space interactively.

The user's answers to the worksheet questions form the skeleton for the user's interaction with the Survey database. That database, in principle, lists the properties of the available software and applications. When addressed with queries derived from the answers developed on the worksheet, it should provide the user either with a selection of applications or display techniques appropriate to the problem, or it should indicate that no such application or technique is known. The latter result should suggest to a developer or researcher that an opportunity exists for novel developments.

2.3.3.1 Using the Results of the Worksheet

Although the Framework is intended to work in tandem with the Survey, the Survey is not essential to the process. Integration of the Framework with the Survey is addressed in Chapter 5. Answering the Framework questions should by itself provide a guide to the best kinds of presentation for the user's problem. Integrating the Framework with the Survey should help the user to find applications that could produce the appropriate kinds of display. In this chapter, we consider mainly the use of the Framework by itself.

If the problem is set in a spatial context, then a display that incorporates the relevant spatial embedding field might well be better than one that displays only the network, as in Figure 2-6a and Figure 2-6b. In Figure 2-6a, the interest is in the internal properties of the network, whereas in 2-6b, the interest is in how the links span the world. In some cases, whether to include an embedding field in the display might be self-evident, but it might be less so in other cases. For example, should a road network be displayed over the landscape? It depends why the road network display is required. If the problem is to determine the driving time on different possible routes from A to B, it probably would be better to display only a graph that shows links on the fastest few routes, whereas if the user's problem is to determine routes suitable for casual tourism, map-based displays including landscape features and tourist attractions (the spatial embedding field) would be better.

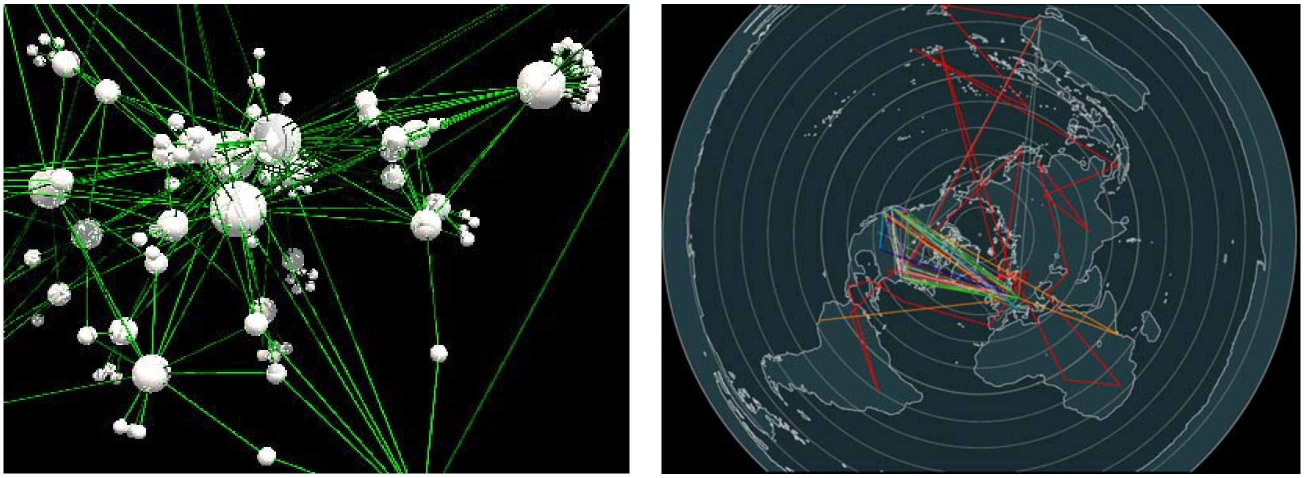


Figure 2-6: Two Views on Parts of the World Wide Web – The left picture (a) shows topical relationships, the right one (b) traffic in a geographic context (an embedding field for the network). (Images are from http://www.visualcomplexity.com/vc/, with permission of the respective authors)

If the problem concerns localized aspects of the network, then the display probably should allow for representations of local detail at the expense of loss of detail for the main body of the network, either by permitting local magnification or by allowing the user interactively to traverse the network in a display that restricts the region shown. Or, if the problem concerns locating aspects of the network that are changing, the display should either permit direct comparison of the network at different times, or should highlight the changed aspects of interest. The interesting changes might be localized, or they might be distributed over the network as a whole. It depends on the user's task, and that dependency should be clarified by following the Framework procedure.

As an example, consider what might be displayed to show the changing links among Al Qaeda cells over time from, say, 1997 to 2007. Over the earlier period, many links would be connected to Afghanistan, but there would probably be only a small number of cells (cells being indicated by sub-nets with higher than normal link density among the people concerned, or “cliques”). For some users and problems, the display might benefit from being overlaid on a world map, whereas for others users and problems the geographic embedding might be irrelevant or distracting.

Changes in the location and structure of cells might be of interest, as, for example, the coalescence of some 9/11 plotters in Hamburg before their move to the USA [11]. If so, then the geographic embedding field might be displayed, but not necessarily in map form. Perhaps locations such as Hamburg and the USA might be treated as nodes in the displayed network, and the move seen as switching the links connecting persons to Hamburg into links connecting the same persons to the USA, with the changed links highlighted in the display. However, after 2003, the interesting changes might include a less localised increase in the numbers of cells in the network, so inclusion of a geographic embedding field in the display might then be less useful, at least for some problems. Alternatively, the increased geographic spread of cells might be important to the user’s task, in which case the embedding field probably should be displayed.

How best to display the network depends on the problem, the user and the viewing situation. A well-designed worksheet should make explicit the issues that the display should address. Evaluation of different display techniques should identify what issues are well addressed by particular kinds of display. The Framework is an attempt to link these two areas of knowledge.

2.4 SUMMARY

The IST-059 Framework for visualisation of networks constitutes a procedure that helps a user to take advantage of taxonomies of data types, network types, network properties, and of display types, to choose a display that most effectively allows the user to see the network properties most relevant to the task at hand. In conjunction with the IST-059 Survey of Network Applications, it should help the user to select appropriate software to generate the displays that the Framework suggests would be most useful.

The Framework is not yet a fully realized tool. The taxonomies of network types and properties are under development, and no software exists to implement the linkage of the procedure to the taxonomies, or from there to the database of available applications and tools. These demand further development before the full power of the Framework can be realized. Nevertheless, by requiring a user to answer directed questions about the task at hand and the network properties that might be relevant, the Framework can even now be useful in helping the user to clarify what kinds of display might help with the task. The user can then manually query the Survey database to seek out software that suits the problem at hand.

2.5 REFERENCES

[1] “Visualisation of Massive Military Datasets: Human Factors, Applications, and Technologies”, *Final Report of IST-013/RTG-002*, RTO-TR-030.

[2] Bouchard, A. and Vernik, R., “Characterisation and Showcasing of Network Visualisation Approaches for Command and Control”, In *Proceedings of Workshop IST-063/RWS-010 Visualising Network Data*, Copenhagen, Denmark, 17-20 October 2006.

[3] Powers, W.T., “Behaviour: The Control of Perception”, (2nd Ed.) New Canaan, Conn: Benchmark Publications 2005 (1st Ed., 1973).

[4] Reenskaug, T., “Thing, Model, View, Editor: An Example from a Planning System”, *Xerox PARC Technical Note*, 12 May 1979.

[5] Reenskaug, T., “Models, Views, Controllers”, *Xerox PARC Technical Note*, 10 December 1979.

[6] Schulz, H.-J., and Schumann, H., “Visualizing Graphs – A Generalized View”, *IEEE Proceedings of the 10th International Conference on Information Visualization (IV06)*, July 2006, London, UK, pp. 166-173.

[7] Smestad, T., “Outline for Concepts and Principles of Effective Visualisation”, *Annex A to RTO-TR-IST-021 – Final Report of IST-021/RTG-007 Multimedia Visualisation of Massive Military Datasets*, 2007.

[8] Shannon, C.E. and Weaver, W., “The Mathematical Theory of Communication”, 1949, Urbana, University of Illinois Press.

[9] Taylor, M.M., “Principal Components Display of ERTS Imagery”, *Proceedings of the Third ERTS Symposium*, Washington, 1973, 1877-1898.

[10] Taylor, M.M., “Principal Components Colour Display of ERTS Imagery”, *Proceedings of the Second Canadian Symposium on Remote Sensing*, 1974, 296-313.

[11] Krebs, E.V., “Mapping Networks of Terrorist Cells”, 2002, *Connections*, 24(3), 43-52.

Chapter 3 – STATE OF THE ART IN NETWORK VISUALISATION AND FUTURE DIRECTIONS

Surveys of existing network visualisation technologies and literature were undertaken in 2006 and completed in early 2007. The technology survey provides a snapshot of the available product capabilities, and the literature survey allows for recommendations for the way ahead for the advancement of network visualisation research and technology in support of military activities. Military users require standards, both for information sharing and to be able to use multiple tools. They also require technical support for the products that they employ, which ultimately leads to a requirement for commercialization. Four areas of focus are identified as essential: sharing data and algorithms across disciplines requires the establishment of standards; representing large, dynamic, and/or uncertain networks is a challenge, as is addressing hardware challenges such as small screen display areas; network visualisation can assist in decision support through exploiting the mathematical properties of networks, creating specialized displays and enabling interactive discovery; and finally, formalizing the evaluation of these methodologies in terms of the human user to prove the effectiveness of a representation or method is essential. With these four areas addressed, the requirements of military users can be met, and the information visualisation community will be able to move closer to achieving their research goals.

3.1 INTRODUCTION

The purpose of this chapter is to provide an overview and review of the state of the art in network presentation technologies, whether they be in production or research stage, and to provide direction on the way forward to advance the field of network visualisation, identifying promising technologies.

First, a survey of existing network presentation technologies was performed, providing a gross view of what capabilities exist and what capabilities haven't reached a stage wherein they may be publicly released. The details of the survey are in Annex F, Section F.1. Second, a literature search was performed, with emphasis on network visualisation research since the year 2000, up to early 2007. This literature search was extended to include "way ahead" papers for the field of information visualisation, since many of the issues related to information visualisation also apply to network visualisation. The literature search provided a gross view of the current fields of interest to researchers in information and network visualisation. The details of the literature search are in Annex F, Section F.2. Taken together, these two surveys provide a snapshot of where we are, and point to where we should be going to advance the field of network visualisation.

It is important to note that neither the literature survey nor the product survey could be exhaustive due to the scope of the problem areas; there is sure to be work of which the survey authors are unaware. Advances since 2007 are not included.

In Section 3.2, the capabilities that are required to advance the state of network visualisation are identified, and the current state of each capability is discussed. In Section 3.3, the military requirements are discussed. A summary of trends and technology gaps is presented in Section 3.4, along with a rough roadmap of the way ahead.

3.2 RESEARCH AREAS

Networks are representations of relationships among entities, and are actively studied in numerous scientific disciplines including physics, biology, computer science, and of course, information visualisation. On the whole, published research in information visualisation is on the rise, as is published research on network visualisation. The portion of these publications interested in node placement algorithms, human aspects and scalability has remained relatively constant. Studies of specific types of networks appeared in 2001, and studies of dynamic networks appeared in 2002; and portion of publications concentrating on these aspects has remained constant since. The InfoVis conference proceedings in particular show that while the total number of papers has remained relatively constant since 2000, the portion of their accepted papers devoted to the visualisation of networks has risen (see Table F-2 in Annex F).

Despite all of this research interest, progression into supported commercial products has been slow. This prompted numerous discussions of trends, challenges and future directions [1]-[6]. In fact, *IEEE Computer Graphics and Applications* has run a regular feature entitled “Visualisation Viewpoints” since 2000, wherein articles along this vein are published [7]-[19]. In 2006, the National Institute of Health (NIH) and National Science Foundation (NSF) in the United States produced a document describing visualisation research challenges [20]. The information provided by these articles applies equally to the sub-discipline of network visualisation.

With the help of these documents and surveys focussed on network visualisation literature [21]-[23], the main areas of research that would contribute to the advancement of the field of network visualisation are identified. The current state of each area, in commercial products and in the research literature, is discussed.

3.2.1 Information Sharing in Multiple Disciplines

If we assert that in order to advance the field of network visualisation we must encourage inter-disciplinary collaboration and cross-pollination, then we must establish standards of discourse that will enable communication [6][10][16][20]. For example, rudimentary to a field of study, one may be studying a “network with nodes and links” or a “graph with vertices and edges”.

3.2.1.1 Data Standards

There is evidence of the desire of researchers and developers to share ideas in the R&D communities, e.g. IBM’s Many Eyes [24], or Visual Complexity [25]. Sharing software tools and algorithms requires that a standard data format be agreed upon, and evaluation will require benchmark data sets that can be used to test all algorithms.

Around 1997, the GML file format was produced to enable data sharing in the graph drawing community [26]. In 2000, the Graph Drawing symposium held a workshop on data exchange formats [27], initiating work on the GraphML format [28][1]. DynetML [29] was developed for use by the social network analysis community for specifying dynamic social networks. There are several data file formats developed by vendors and developers, including the proprietary i2 Analyst’s Notebook [30] file format (ANB), which is used by NATO [31] and the USA military [32].

The GraphML data format standard is not finalized. Thus far, a standard data format has not been adopted by the scientific community. In the product survey, of the products using an open format, 39% used GraphML.

---

[1] Other file formats include XGMML (http://www.cs.rpi.edu/~puninj/XGMML/), GXL (http://www.gupro.de/GXL/), and SVG (http://www.w3.org/Graphics/SVG/), but these were not seen in the product survey.

3.2.1.2 Benchmark Data

Benchmark data is needed to test new algorithms, and to evaluate and compare tools [33]. The following data sets were discovered in this review:

- The Information Visualisation Benchmark Repository [34]:
    - Pair Wise Comparison of Trees (InfoVis 2003 contest);
    - The History of InfoVis (InfoVis 2004 contest);
    - Technological trends in the United States (InfoVis 2005 contest);
    - Microdata sample from 2002 USA Census (InfoVis 2006 contest); and
    - A tale of Alterwood (VAST 2006 contest).
- The 1998 and 1999 DARPA Intrusion Detection Evaluation Data Sets and the 2000 DARPA Intrusion Detection Scenario Specific Data Sets [35]. These data sets contain simulated computer network traffic, some having labelled attacks and some having no attacks.
- The Internet Mapping Project [36]. This is a large computer network data set, which shows linkages between nodes on the Internet from 2006.
- Di Battista et al. [37] generated a benchmark data set of 11,582 graphs ranging from 10 to 100 vertices, however this data is no longer available at the cited URL.
- The Knowledge Discovery and Data mining (KDD) Cup 2003: “This KDD Cup is based on a very large archive of research papers that provides an unusually comprehensive snapshot of a particular social network in action” [38].
- Benchmark data could also be simulated by graph generators, a survey of which is given in [39].

3.2.1.3 Taxonomies

In order to classify information visualisation techniques, taxonomies have been developed for data types [40] and for tasks [40]-[42]. One taxonomy of the visualisation process itself, from data collection to presentation, was used to classify and identify similarities in several visualisation techniques [43].

Pattison et al. [44] created a taxonomy of layout strategies for attributed graphs. Schulz and Schumann [45] present a taxonomy of network representations along with user, data and aesthetic constraints to build a process-based framework for network visualisation. Lee et al. [46] present a task taxonomy specifically for graphs, and use their taxonomy to compare and classify five graph visualisation tools. Other taxonomies and frameworks have been presented for information visualisation, e.g. [47]-[57], and for network visualisation, e.g. [39], [58], [59].

In 1988, Tamassia et al. [60] recommended future work in developing a framework that would enable “a parametric algorithm that can be interactively tailored to specific classes of diagrams, graphic standards, aesthetics, and constraints.” Twenty years later, a complete and common taxonomy of network tasks or network layouts has not yet been adopted, let alone a general graph drawing framework.

3.2.1.4 Software Languages and Architectures

Modular software architectures can allow easier sharing of network layout algorithms. Modular software architectures are found, for example, in GVF [61], IVC [62], Tulip [63], and OGDF (formerly AGD) [64]. In these software architectures, users may contribute modules that they have developed for graph layouts.

These are programmed in Java, and C++. The programming environments MATLAB [65] and Mathematica [66] do not explicitly display the network, but offer functions to allow graph and animation coding to be easily developed.

There were 20 different languages used in the products discovered in the product survey, almost 50% of which are programmed in Java and 20% in C++. This variety of languages makes it difficult to share code. The NetworkX [67] package makes use of Python, a flexible multiplatform language that can be used to integrate several languages into one application. The same site houses PyGraphviz, a Python interface to the Graphviz graph layout and visualisation package [68].

3.2.2 Network Representations

For each application area, there is one or more type of network under examination and there may be several different tasks that the user may wish to undertake. Blythe et al. [69] show that for social network analysis (SNA) the layout chosen for a network has a significant effect on the inferences drawn by a user. The most effective representation of the network depends on the task, and possibly other factors. In this section, we discuss what has been done in the commercial and academic sectors to handle large networks. Further, we discuss the state of the art in representing error and uncertainty on a network, and discuss special considerations for hardware. Finally, we discuss temporal representations for change detection and trend detection.

3.2.2.1 Large Networks

The increasing volume of data to which we have access gives rise to the problem of effectively dealing with large data sets [11][13][17][70]. For example, a moderately-sized “class B” computer network may contain up to 65,536 nodes, and a “class A” computer network may contain up to 16,777,316 nodes.

Getting useful information out of a large network dataset requires simplification of the data and/or user interaction of some form, e.g. zoom, pan, distortion, filtering, clustering – all the while maintaining a global context. Simple zoom and pan techniques are often applied, e.g. Google Earth [71], but they do not, in themselves, give the user context, e.g. where they are on the globe, when they zoom in to fine detail.

In the product survey, 79% of the products did not indicate their scalability, 16% of the products claimed unlimited scalability, and the remaining 5% of the products indicated being limited to less than 100,000 nodes. Note that this field in the survey does not give an indication of how aesthetically pleasing a layout is for large networks, nor how usable the visualisation is for a particular task, only whether the software is capable of processing large networks in reasonable time scales. It remains inconclusive whether there is a commercial solution to the large network problem.

3.2.2.1.1 Maintaining Global Context

Because of the large size, it is difficult to present local detail while also giving global context to the user. The “overview+detail” technique uses two windows; one shows a reference map of the full graph, and marking a region shows details of the region in another window.

Distortion techniques (a.k.a. “focus+context”) allow one to enlarge regions of interest while using just one window. These have been used for several years in network visualisation, including the graphical fisheye view [72], and other lenses such as the “bring neighbours” lens [73] and the hyperbolic tree [74]. Variations have been developed, such as the 3D hyperbolic tree [75]. Products that have distortion capabilities include aiSee [76], H3Viewer [77] and the InfoVis Toolkit [78].

Baudisch et al. [79] found that using focus+context techniques allowed users to complete their tasks faster than with either overview+detail or zooming and panning techniques. They also showed that focus+context techniques reduced operator error. For viewing large interfaces on small screens, Gutwin [80] found that the fisheye view allowed user to carry out Web navigation tasks more efficiently than pan and zoom techniques. On the other hand, Hollands et al. [81].showed that for complex tasks, the fisheye view can be confusing.

3.2.2.1.2 Network Simplification Techniques

One can approach the large graph problem by improving the layout of the edges in existing representations. For example, Gansner and Koren [82] build on the circular layout by implementing edge bundling and allowing links to be shown on the exterior of the circle. Holten [83] experiments with bundling edges in large hierarchical graphs and applies alpha blending, which emphasizes short curves by drawing long curves at a lower opacity than short curves.

The large graph problem can also be addressed through data reduction and node aggregation (clustering) techniques. In evaluating the technical aspects of these techniques, one must ensure that data reduction does not delete important data [12], or obscure important details. For example, computer networks are generally sub-divided into smaller and more manageable sub-networks. This lends itself naturally to visualising the network based on collapsing these assigned sub-network regions. The “opening and closing” of clusters is presented in [84]. The capability to expand and collapse groups of nodes is available in the Jgraph open source Java graph library [85]. If this technique is used, however, there is an additional problem of being unable to detect multiple changes to a collapsed group, e.g. in a monitoring activity, after a node has turned red due to a problem in one of its sub-nodes, it remains red when problems arise in other sub-nodes.

The sparse properties of scale-free networks are exploited in [86] to simplify a graph while maintaining the underlying graph patterns. This method clusters nodes into a single representative node based on shortest paths. Bjørke aggregates nodes and links into “hyper-nodes” and “hyper-links” by reordering the adjacency matrix of the network, thereby generating hierarchies of hyper-networks [87]. Motifs are recurring structures within a network. Substituting a motif in a network can simplify its display for the human user [88][89].

Moody [90] pointed out in that large diagrams should be divided into chunks such that their number does not exceed Miller’s magic number, 7±2, referencing his earlier work that showed that modularising information system diagrams improved end-user understanding by more than 50%.

3.2.2.1.3 Effective Use of Screen Space

Large hierarchies can be displayed by showing each level of the hierarchy in a plane of its own [91], or by the RINGS technique [92], where each singly-linked sub-graph is shown as a circle, with its child sub-graphs contained within it, and their children within them, and so on.

Partitioning the network into different layout types may help with using screen space effectively. In their 1988 paper [60], Tamassia et al. recommend that future research include “devising a layout strategy that allows the use of different graphic standards for different parts of the diagram”. A partitioning algorithm is shown in [93]. In [94], a metric of networks is exploited to identify links that, if broken, will break down the network into smaller components that are easier to comprehend. The algorithm presented in Dwyer et al. [95]

automatically identifies hierarchical information and represents it using a directed graph algorithm, and draws the non-hierarchical information with a non-hierarchical undirected layout algorithm. Vandenberghe's layout [96] places leaf nodes in a rectangle to improve the use of screen space.

3.2.2.1.4 Algorithm Speed

Walshaw [97] groups adjacent nodes to define a new graph and the process is repeated recursively, resulting in a set of increasingly coarse graphs. The coarsest graph is first laid out using a force-directed method, then the next level is added and the layout refined, and so on until the original graph is shown. This is shown to accelerate the process of laying out a very large graph, although 225,000 nodes still requires 5-7 minutes. It purports to result in a drawing with "a more global quality".

Chan et al. [98] find that in power-law network topologies, the highly-connected nodes are the most influential in determining the final structure in a force-directed layout. By progressively laying out nodes of highest out-degree, they improve the speed of the process.

Vandenberghe [96] presents an algorithm that places network nodes significantly faster than force-directed methods. Force-directed methods require several iterations to reach an equilibrium state, whereas Vandenberghe's Voting algorithm requires a single pass.

3.2.2.1.5 Non-Node-Link Representations

Node-link diagrams may not be the optimal way to present some types of networks or network information. The InfoZoom tool was found to perform quite well in the 2003 InfoVis contest [33], using tables instead of node-link representations. Ghoniem et al. [99] find that matrix-based representations are more suitable for large or dense graphs than node-link diagrams, due to occlusion problems in node-link diagrams. They propose increased exploitation of this method of presenting large networks. MatrixExplorer [100] is a system that shows synchronized matrix and node-link representations of a network (a coordinated multiple view, as in Section 3.2.3.2), and also incorporates other important concepts such as user interaction and overview maps.

3.2.2.2 Representing Uncertainty and Unknowns

Uncertainty may be defined as the difference between the reality and a perception of that reality. Uncertainty is inherent in measured or observed data, and must be communicated to the user [9][11]. Further, a means of reporting that a data element is unknown or unreliable must be developed. Uncertainty does not always need to be represented on the screen to be understood by the user; other ways such as proper training may sometimes be sufficient.

Howes [101] reviewed different techniques to represent uncertainty and complex information, and how people make decisions under conditions of uncertainty. A main conclusion of this study is that technologies to represent uncertainty already exist, although the efficacy of these representations when presented to the user in real world decision spaces has not been thoroughly evaluated. The nature of the uncertainty (e.g. identification, accuracy and reliability of the source, data gaps) has certainly an impact on how the data should be represented and how the human copes with the uncertain data. In 2005, the Visualisation Network of Experts workshop produced a presentation of how errors and uncertainty can be represented for networks [102].

In the context of a network, one needs to be aware that the absence of a node or a link is not an indication that the node or link does not exist. Users typically have to deal with networks that are gross estimations of reality, especially for non-physical networks such as social networks. The cost of acquiring complete and accurate data may be extremely high in some cases or not possible in other cases, and the user may have to make decisions knowing that the data is incomplete and uncertain.

In the collection process, data normalisation can be an issue. For example, the location of DRDC Ottawa may correctly be referred to as "Shirley's Bay", "Kanata", or "Ottawa", and any of these may be misspelled. The D-Dupe application [103] provides an interactive capability to reconcile duplicate nodes, i.e. nodes that are the same entity with different labels.

3.2.2.3 Special Displays and Hardware

Of the 139 products in the survey, only 5 required specialised hardware: 4 required a 3D graphics accelerator and 1 required an electronic whiteboard with click and drag. Eick [14] points out that as new hardware technologies emerge, we must re-think the way visualisations are done. For example, the advent of portable devices such as the PDA requires a re-engineering for visualising information on small displays. One could envision the case of a computer network administrator away from their desk, and diagnosing a trouble call using a hand-held device. Likewise, large displays such as wall displays and other large screens remove size restrictions, however the user's processing limitations must be taken into account [13]. TTCP C3I TP2 is studying the use of large-screen displays [104].

Other interesting emerging technologies will allow for creative information presentation: the 3-dimensional cylindrical television [105] may offer unique opportunities, and Microsoft's Surface [106] provides an interactive large-space display area.

Johnson [11] suggests the potential use of specialised hardware such as the graphics processing unit (GPU), multiple graphics cards, and distributed grid-based computing. Frishman [107] implements a force-directed algorithm on the GPU by partitioning a large problem into smaller, similarly-sized problems.

3.2.2.4 Temporal Representations

When data is time-dependent, a visualisation should draw the user's attention to a trend or a change [11][14][17]. A recent showcase of dynamic visualisation methods is given in the Competition on Visualising Network Dynamics [108].

When the user's task is to detect change in a network, it is imperative that the user be able to relate what they see at time $t+\Delta t$ to what they saw at time t. This is known in the literature as *dynamic stability*, or *preserving the mental map*. When the task is trend identification, the change occurs over a series of time steps and preserving the mental map is equally important. This can be difficult to accomplish, especially for large networks. With many tools, the addition or deletion of a node or link in a network can result in a substantial change to the layout, causing difficulties in understanding how the new network relates to the old.

3.2.2.4.1 Change Detection

A common approach to preserving the mental map of a network consists of animating the movement of the nodes from the original position to the position in the next time step. Kapler and Wright [109] developed a commercial product called GeoTime which implements a space-time cube for network visualisations.

The software presents a 3D geospatial display where the third dimension is time. The ground represents the current time and the time increases as one moves away from the ground. Animation techniques are also available in, e.g. aiSee [76], ILOG Jviews [110], Nevron Diagram [111] and yFiles [112].

The animation approach is adequate for small networks but becomes ineffective in large networks, where the movement of a large number of nodes at the same time is overwhelming to the user. Some products (e.g. Tom Sawyer [113], yFiles [112]) use the "incremental layout" rendering technique to preserve the mental map. This technique minimizes the spatial movement of nodes when rendering a new layout. While this new layout may not be the optimal layout, it is close to the original one, and the user can activate the incremental layout function repeatedly until a stable layout is reached, with minor movements at each iteration. Such a technique minimizes the cognitive overload of rebuilding the mental map after each rendering of a layout.

Another technique used to preserve the mental map in dynamic networks consists of locking the nodes under investigation to ensure that they don't move out of the working area when activating a new layout [114]. When a rendering function is activated the locked nodes will not move and the other nodes position themselves around the locked nodes.

In [115], the network is displayed on planes, with each plane representing a time step. This work is extended in [116], where instead of showing the full network at each layer, only the difference between the network at time t and time t + Δt is shown; stacking all "difference layers" shows the full network. Drawing of edges not only between nodes but also between clusters is discussed in [117][118], concluding that this type of layout provides dynamic stability. Similar work is presented in [119], where a cluster is represented by an icon showing the properties of the cluster.

3.2.2.4.2 Trend Detection

The "network movie" for trend detection is available in three products identified in the technology survey. GraphAEL [120], a general graph drawing tool, offers smooth animation and uses fading when new nodes enter/exit. For social networks, SoNIA [114][121] can animate transitions between network configurations and TeCFlow [122] allows the user to choose how much historical information to include, with older information shown faded in the display. These software packages are research tools, freely available to the community.

In [114], it is found that static flip books, where node position remains constant but edges cumulate over time, are particularly useful in contexts where relations are sparse. Network movies, where nodes move as a function of changes in relations, are more appropriate for more connected networks.

In [123], minimum spanning trees are compared to pathfinder networks for visualising evolving networks. It is found that pathfinder networks are better suited to dynamic network visualisation because it shows both local and global structural evolution.

3.2.3 Decision Support

The purpose of visualising information is to generate insight into the current situation. Insight may be defined as the perception, comprehension, and projection of the inner nature of things. The user should be able to draw conclusions from what they observe, and from there make better-informed decisions [17]. Awareness of the current situation can often be presented more readily via visual means, if the data set to be absorbed is large.

3.2.3.1 Mathematical Properties of Networks

Network science has only recently begun to be established as a generic field of study borne from the application areas in which networks frequently appear. The theoretical aspects of network analysis are well established, regardless of terminology differences among domains of interest. In [124], Börner provides a chapter that brings diverse network analysis terminology, primarily from social network analysis and physics, together into one document.

Comprehension of properties of the network can be supported by visualisation by providing context. For example, when Scipio Optipath [125] determines the optimal route through a road network, it is viewed overlaid on the underlying network to provide context.

We can provide further understanding of the network by displaying measured properties of the network. For example, Auber et al. [94] describe a metric that assists in identifying the weakest edges in a small world network. Perer and Shneiderman implement user selection of network measurements to colour nodes according to their measured values [126].

The discovery of inter-node dependencies provides information in support of decision-makers, for example, the propagation of effects within social networks, or propagation of “asset values” in a computer network. An algorithm to determine the relative importance of a network node, in terms of the dependencies of other nodes upon it, is presented in [127]. The discovery of dependencies among apparently disparate networks will also aid the decision-maker, e.g. computer networks depend on hydro networks.

3.2.3.2 Multiple Data Views

Coordinated and multiple views (CMV) are displays that contain multiple, linked representations of the same data set, such that interaction with one view leads to changes in all views. These have been suggested as a means to support the processes of exploration and discovery [8][17]. Some studies have been conducted in this area (e.g. [128]-[130]); in fact there has been an annual conference on the topic since 2003 (Coordinated and Multiple Views in Exploratory Visualisation). CMV are implemented in Pattison et al. [44], but this technology has not been exploited by vendors. This may be due to data set size restrictions; CMV is limited to about 10^5 records due to the computational power required and 10^7 records due to screen size [131].

Other representations have been suggested for enhancing the ability to deliver knowledge, such as visualising multiple data sets on the same surface [132] or otherwise simultaneously [11]. This is important for cases where networks exist in layers or where networks share nodes. For example in a computer network, there is a physical layer and an application layer [133]. If a node is removed from the physical layer, the application layer is also affected. The concept of “logical overlays” was addressed in 1995 for computer networks [134], wherein the authors identified the importance of being able to visualise the relationship between the physical and logical layers. These interdependencies can also be viewed by using “semantic substrates” [135], where nodes for each layer are shown on different areas of the screen. Links between the areas clearly show dependencies between layers. In [136], schemes for laying out networks that share nodes are presented: the aggregate view (simultaneous), merged view (multiple layers) and split view (side-by-side) models.

3.2.3.3 Interactive Discovery

Networks can be large, and complex in node and link attributes and their relationships. In these cases, a static display will not be sufficient to relay enough information to the user to make a decision. Interactive capabilities allow the user to steer the visualisation toward regions of interest [12][137] to obtain greater detail where it is needed, through drill-down or the application of lenses. The user may wish to choose the most appropriate method of decomposing a complicated graph, depending on their region of interest [134], hence requiring the ability to change the layout algorithm or parameters. The user may wish to filter the data presented, or colour nodes or links based on some attribute or statistical measurement [126]. Allowing the user to explore the network gives greater opportunities for discovery.

Modular architectures can also be used as part of the discovery process, where the user chooses one or more algorithms to operate on the data. In [44], Pattison et al. present a software environment where the user builds their tailored view of the network data. The resulting views may be as interactive as desired, and may contain multiple coordinated views of the data. This is in essence a “visualisation-building environment”.

3.2.3.4 Prediction

Johnson [11] asserts that to allow the user to perform what-if scenarios, or steer computations on-the-fly, environments must be developed that allow the user to simultaneously model, simulate and visualise. Visualisation could aid in the “what-if” scenarios by allowing the user to quickly assess the behaviour of the network under changes to a node or link. For example, removing a node in a social network may result in the network fragmenting; removing a link in a computer network may result in an impact on business operations. No literature on this topic was discovered.

3.2.4 Evaluation of Network Visualisations

The evaluation of information visualisation systems has been cited as a high-priority challenge in several documents, e.g. [17], [70]. Evaluating a visualisation technique or system is required to prove its effectiveness, both alone and in comparison to other methods. Moody [90] notes that some information may be better represented in textual form. Human-computer interaction (HCI) studies these aspects and has been around for some time, but has not been well integrated with the information visualisation field [11]. Tory and Möller review known methodology for human factors research and the state of human factors research in visualisation, describing several promising areas for future research [55]. An annual conference was initiated in 2006 to address the evaluation of information visualisation (BELIV: BEyond time and errors: novel evaLuation methods for Information Visualisation).

3.2.4.1 Cognitive Evaluations

Evaluations of visualisation techniques necessarily include some technical aspects, such as computational speed and use of screen space, however the bulk of the evaluations that need to be done are human-centered, i.e. meaning is conveyed to the human in adequate time and with adequate accuracy.

Formal laboratory user studies are the standard in evaluating information visualisation systems, requiring substantial time and resources. Purchase et al. [138] performed a study of the understandability of graphs based on the technical aesthetics of arc bends, arc crossings and symmetry, finding that both bends and crossings should be minimised to increase understanding, while symmetry had an inconclusive effect. Ghoniem [99] bases an evaluation methodology on seven generic graph analysis user tasks to compare matrix-based representations to node-link diagrams, finding that node-link diagrams are favoured only for path-finding tasks. This paper also provides a review of the evaluation techniques up to that point.

Some experts are not convinced that formal user studies are always appropriate [18][33]. The level of knowledge of the user study participants may not reflect the knowledge level of the intended end-user [17].

By forming a tiger team of experts in HCI, visualisation, graphic design and end-user tasks, an expert evaluation can be performed on a far shorter timeline and with fewer subjects, which can be very useful in preliminary evaluations. This, however, should not replace user studies [18]. In [139], Xu and Chen evaluate their CrimeNet system by comparing the results of the system with the results obtained by human experts. In the Imago environment [57], a semantic model based on the RM-Vis framework is queried to give candidate views to the expert user, displaying many potential views of the same data. Expert users provide assessments of the effectiveness of each view, given the set of conditions (user tasks or goals, and data types).

For complex systems that must provide situational awareness and decision support via exploration, approaching the evaluation using tasks may not be appropriate [140]. Multidimensional In-depth Long-term Case studies (MILCs) are an evaluation method developed to support “creativity support tools” (tools for long-term exploratory tasks) [135], which avoids defining user tasks.

3.2.4.2 Aesthetics

Although graph aesthetics and readability were investigated as early as 1988 [60], Hibbard [13] and Chen [17] noted the lack of study of what makes a visualisation aesthetically pleasing in 2004 and 2005, respectively. In 2005, Keefe [15] recommended that artists be embedded in the design process when developing visualisations. Moody [90] argued in 2007 that the decision of which layout to use should be based on evidence about cognitive effectiveness rather than aesthetics. However, in the same year, Cawthon and Vande Moere [141] published a study on the effect of aesthetics on usability, with a focus on tree structures, concluding that the most aesthetically pleasing visualisation techniques have a lower rate of task abandonment, and enable the user to provide more correct responses in less time.

3.2.4.3 Guidelines for Drawing Graphs

Based on some of the work that has been done in cognitive and perceptual psychology, Moody has presented 9 principles for producing effective diagrams [90]. Huang et al. [142] also list a set of rules, derived from qualitative results.

3.3 MILITARY NEEDS FOR NETWORK VISUALISATION

Military application areas for network visualisation include computer network defence, net-centric warfare, terrorist networks, and the spread of infectious agents. These application areas require decision support. Because the military user requires trust in the results delivered, the uncertainties and the logic used to arrive at a conclusion must be readily presented. For coalition information sharing, militaries need a standard data format so that users can analyse data with different software. Technical support must be provided to the users, as military organizations typically do not necessarily have access to experts to assist them with unsupported code, which ultimately leads to a requirement for commercialization.

These needs will require defence scientific staff to push industry and academia toward developing and using a standard data format. They must also drive academia to perform evaluations, which will encourage industry to develop and support a usable product.

3.4 THE WAY AHEAD

This section bears many similarities to the way ahead presented by Thomas and Cook [19] and by the USA National Institute of Health [20]. To maximize the potential for creating successful network visualisations, we need to coordinate our activities in the various application domains and begin to share our knowledge of representing networks. To enable this sharing, we must all adopt a set of common standards, such as those suggested by the IST-059 Framework (Chapter 2 and Annex B). Then we may begin to recognize cross-domain solutions and continue to develop representations of network data that will not only enhance our ability to absorb data efficiently, but present the data in such a way as to enable us to see something we would not have otherwise seen. Evaluations of network representations must become a standard practice, to prove effectiveness to colleagues and to assist in convincing venture capitalists that a technique is worthy of commercialization (known as “crossing the chasm” [33]).

3.4.1 Sharing

Overall, the field of information visualisation lacks standards. To enable collaboration between multidisciplinary experts, a common language of discourse is needed. Work has been done in the development of a theory of network visualisation, but a single framework has to be accepted by the community for progress to be made. We have not formally accepted a standard data format, and we lack a means of standardizing user input.

Of the surveyed network visualisation products (toolkits), 40% use Java, which operates on many platforms, but is not universally used in research and academia.

The way forward requires that we need to:

1) Agree on and adopt a standard data format. The GraphML format should be evaluated and finalized or discarded.
2) Build a standard network data repository. Several data repositories exist, however they are very application-specific. To maximize usability, a generic set of data could be developed that categorizes and anonymizes data that is contributed to a repository. An initiative such as IBM’s Many Eyes [24] constitutes a good start to create a dataset repository for information visualisation. The uploaded datasets have to comply to a format predefined by IBM. However, there is currently no initiative specific to network visualisation.
3) Develop generic taxonomies for tasks, layouts, and aesthetics that are extensible for characterizing more specific, domain-dependent, tasks.
4) Develop a framework that will allow researchers and users to generalize the proper display type for the given type of data and the task. Although designing based on task specification does not necessarily allow for creativity, which is needed for the discovery process [135], nevertheless it is still necessary to verify what visualisation technique works for the tasks that are well-defined.
5) Assemble all of the modules in a common repository for all researchers to access. The InfoVis Cyber-infrastructure is a step in the right direction, however it does not accommodate a researcher’s preferred programming language. To address this issue, one might consider using the Python programming language [143] as a “glue” to patch together modules programmed in the language of a researcher’s choosing, e.g. Java, C++, or MATLAB.

To achieve maximum buy-in from the graph drawing and network/information visualisation communities, a trusted and diverse network of experts must be assembled to collaborate on these standards.

3.4.2 Representations

Many representations of networks have been developed; often the representation used is tailored to the application. For large networks, overview and distortion techniques have been used to provide global context, and clustering and other means have been used to show the user the underlying structure of the network. Several techniques have been developed to handle the rendering speed. Still, many large network issues remain:

1) A comprehensive evaluation of existing methods must be performed to evaluate their usefulness. It may be that some data is better shown in tabular format.
2) How can one visualise a large data set on a small device?
3) When collapsing a cluster into a single node, how do we retain awareness of change within the collapsed cluster? Dashboards are commercially available [144] and may provide the global context that data reduction and node aggregation can obscure.
4) More work should be done to determine the human's comprehension of changes to a large, dynamic network. Preserving the mental map must be a primary consideration.

While techniques for representing uncertainty exist, their efficacy has not been formally evaluated. The problem of representing networks on very small or very large screens has not been addressed. We have hardly scratched the surface of using our other, non-visual perceptions to interpret data, such as sound, smell, touch, and taste. Multiple views of multiple modalities are explored in [145], where bar charts are presented with simultaneous visual and auditory aspects. Some work has been done in the use of immersive environments for computer network security [146][147].

Visualising temporal changes in networks, for either change detection or trend detection, has been well-examined, but again, the efficacy for the end-user has not been formally evaluated. In terms of technical performance, procedural generation methods may be worth investigating. For example, in the first-person shooter game ".kkrieger" [148], all assets used in game play are produced during the loading phase and animation takes place on-the-fly, resulting in a 97 kB application, which if stored conventionally would require 200 – 300 MB.

3.4.3 Decision Support

Network science has begun to be established as a field in its own right, as has visual analytics. Both of these contribute to the use of network visualisation for decision support applications.

The meaning of statistical network properties must be re-evaluated in the context of each application area. Social network analysis makes use of many mathematical properties; these can be transferred to the other application areas to convey similar meanings. Visualising these properties on a network has been done, and should be continue to be investigated.

Coordinated multiple views have been proven to be effective in conveying more information than one view alone, especially for showing inter-node dependencies. It is possible that this technology has not been transferred to a commercial product due to an environment needing to be application-specific. A system like that in [44], together with a modular and easily tailored environment, may address this need. Starlight [149] and Jigsaw [150] are examples of research prototypes integrating multiple coordinated views of network visualisations for intelligence.

More emphasis is needed on the interactive aspects of the displays. It has been suggested that 3D gaming environments can provide enhanced functionality [8][151]; this could be applicable in an interactive discovery process, however it is unclear how one would want to explore a network in this fashion.

The decision support functions of what-if scenarios, prediction and hypothesis testing have not been addressed in the literature or in technology.

Finally, as some data is hard to get, we have to also determine the value of data in achieving a task: how important is it to have this data in order to draw a conclusion? Whether this has been addressed in another field is not known.

3.4.4 Evaluation

We need to evaluate what currently works, and what doesn't. Whether visualisation can convey information better than other methods needs to be investigated. Evaluations of how the network visualisation techniques are perceived by the human user are required in order to make decisions about the way forward.

Determining what visualisation technique works best for what data characteristics and what task is challenging for many reasons, not the least of which is that we have not determined which evaluation technique should be applied, given the exploratory nature of some domains. A standardized evaluation framework should be implemented; one presented for the evaluation of Command and Control technologies [152] could be modified and applied specifically to network visualisation approaches.

The evaluation step is a critical part of the development process, because it tells researchers whether they're going in the right direction, or if they should back up and try another approach. Application-domain researchers therefore can no longer work in isolation as they often have in the past; psychology expertise is required to understand a system that includes the computer *and* the human. Johnson suggests that a study of the biophysics and psychophysics of the visual system to guide visualisation methodologies may be beneficial [11].

Aesthetics have been shown to be important to the human user, and so we should increase our interaction with artists and graphic designers.

Trust is an issue that has not been addressed for information visualisation.

3.5 SUMMARY

In this chapter, four areas of focus were identified as being required to advance the network visualisation field.

- **Information Sharing Support**, which includes theory, standards, and software, is needed to allow researchers from diverse application domains to work together. Working together across disciplines will enhance creativity.
- **Network Representations** must be improved to provide satisfactory presentations of large and/or dynamic networks, along with an indication of uncertainty. They should be adaptable to specialized hardware.
- **Decision Support** is often the end goal of displaying the data to the user. The unique mathematical properties of networks can be exploited to assist the user in accomplishing this task. Prediction of future network behaviour is an unaddressed research area.

- **Evaluation** must be integrated into the research process. If a method is to be accepted by the community, good science requires evidence that the method satisfies a human user. If a method is to be transitioned into a commercial product, industry requires some assurance of its efficacy.

Network visualisation is a fairly new discipline and its foundation is still to be defined and accepted by the scientific community. Advances in the domain of information visualisation in term of standards, representations, and evaluations will necessarily benefit network visualisation. Building on the good work already done and standardizing the evaluation process will help focus our efforts. We may one day even reach Tomassia et al.'s utopian vision of "a parametric algorithm that can be interactively tailored to specific classes of diagrams, graphic standards, aesthetics, and constraints" [60].

3.6 REFERENCES

[1] Hascoët-Zizi, M., Ahlberg, C., Korfhage, R., Plaisant, C., Chalmers, M. and Rao, R., (1996), "Where is Information Visualization Technology Going?", In *UIST '96: Proceedings of the 9th Annual ACM Symposium on User Interface Software and Technology*, pp. 75-77, New York, NY, USA: ACM Press.

[2] Tollis, I.G., (1996), "Graph Drawing and Information Visualization", *ACM Computing Surveys*, 28(4es), 19.

[3] Hibbard, B., (1999), "Top Ten Visualization Problems", *SIGGRAPH Computer Graphics*, 33(2), 21-22.

[4] Jaech, J., North, S.C., Peery, M., Schroeder, W. and Thomas, J., (2003), "The Visualization Market: Open Source vs. Commercial Approaches", In *Proceedings of the 14th IEEE Visualization 2003 (Vis'03)*, pp. 21-24, Washington, DC, USA: IEEE Computer Society.

[5] Rhyne, T.-M., Hibbard, B., Johnson, C., Chen, C. and Eick, S., (2004), "Panel 1: Can We Determine the Top Unresolved Problems of Visualization?", In *VIS '04: Proceedings of the Conference on Visualization*, pp. 563-566, Washington, DC, USA: IEEE Computer Society.

[6] Laramee, R.S. and Kosara, R., (2007), "Future Challenges and Unsolved Problems (in Human-Centered Visualization)", In A. Kerren, A. Ebert, and J. Meyer (Eds.), *Human-Centered Visualization Environments – Springer Lecture Notes in Computer Science (LNCS) Tutorial Volume 4417,* pp. 231-254, Dagstuhl Castle, Germany: Springer Verlag.

[7] Rhyne, T.-M., (2000), "Scientific Visualization in the Next Millennium", *IEEE Computer Graphics and Applications*, 20(1), 20-21.

[8] Shneiderman, B., (2003), "Why Not Make Interfaces Better than 3D Reality?", *IEEE Computer Graphics and Applications*, 23(6), 12-15.

[9] Johnson, C.R. and Sanderson, A.R., (2003), "A Next Step: Visualizing Errors and Uncertainty", *IEEE Computer Graphics and Applications*, 23(5), 6-10.

[10] Rhyne, T.-M., (2003), "Does the Difference between Information and Scientific Visualization Really Matter?", *IEEE Computer Graphics and Applications*, 23(3), 6-8.

[11] Johnson, C., (2004), "Top Scientific Visualization Research Problems", *IEEE Computer Graphics and Applications*, 24(4), 13-17.

[12] Ward, M., (2004), "Finding Needles in Large-Scale Multivariate Data Haystacks", *IEEE Computer Graphics and Applications*, 24(5), 16-19.

[13] Hibbard, B., (2004), "The Top Five Problems that Motivated My Work", *IEEE Computer Graphics and Applications*, 24(6), 9-13.

[14] Eick, S.G., (2005), "Information Visualization at 10", *IEEE Computer Graphics and Applications*, 25(1), 12-14.

[15] Keefe, D.F., Karelitz, D.B., Vote, E.L. and Laidlaw, D.H., (2005), "Artistic Collaboration in Designing VR Visualizations", *IEEE Computer Graphics and Applications*, 25(2), 18-23.

[16] Duke, D.J., Brodlie, K.W., Duce, D.A. and Herman, I., (2005), "Do You See What I Mean?", *IEEE Computer Graphics and Applications*, 25(3), 6-9.

[17] Chen, C., (2005), "Top 10 Unsolved Information Visualization Problems", *IEEE Computer Graphics and Applications*, 25(4), 12-16.

[18] Tory, M. and Moller, T., (2005), "Evaluating Visualizations: Do Expert Reviews Work?", *IEEE Computer Graphics and Applications*, 25(5), 8-11.

[19] Thomas, J.J. and Cook, K.A., (2006), "A Visual Analytics Agenda", *IEEE Computer Graphics and Applications*, 26(1), 10-13.

[20] Johnson, C.R., Moorhead, R., Munzner, T., Pfister, H., Rheingans, P. and Yoo, T.S., (Eds.), (2006), "NIH-NSF Visualization Research Challenges Report", 1st Ed, Los Alamitos, CA, USA: IEEE Press.

[21] Di Battista, G., Eades, P., Tamassia, R. and Tollis, I.G., (1994), "Algorithms for Drawing Graphs: An Annotated Bibliography", *Computational Geometry Theory and Applications*, 4(5), 235-282.

[22] Herman, I., Melançon, G. and Marshall, M.S., (2000), "Graph Visualization and Navigation in Information Visualization: A Survey", *IEEE Transactions on Visualization and Computer Graphics*, 6(1), 24-43.

[23] Kershenbaum, A. and Murray, K., (2005), "Visualization of Network Structures", *Journal of Computing Sciences in Colleges*, 21(2), 59-71.

[24] IBM, (2008), "Many Eyes" (online), http://services.alphaworks.ibm.com/manyeyes (Access Date: 15 January 2008).

[25] Lima, M., (2008), "visualcomplexity.com | A Visual Exploration on Mapping Complex Networks" (online), http://www.visualcomplexity.com (Access Date: 15 January 2008).

[26] Himsolt, M., (1997), "The GML File Format" (online), Universitt Passau, http://www.infosun.fim.uni-passau.de/Graphlet/GML (Access Date: 23 January 2008).

[27] Brandes, U., Marshall, M.S. and North, S.C., (2001), "Graph Data Format Workshop Report", In *GD '00: Proceedings of the 8th International Symposium on Graph Drawing*, pp. 407-409, London, UK: Springer Verlag.

[28] Brandes, U., Eiglsperger, M., Herman, I., Himsolt, M. and Marshall, M., (2001), "GraphML Progress Report, Structural Layer Proposal", In Mutzel, P., Junger, M., and Leipert, S., (Eds.), In *Graph Drawing – 9th International Symposium, GD 2001 Vienna Austria*, pp. 501-512, Heidelberg: Springer Verlag.

[29] Tsvetovat, M., Reminga, J. and Carley, K., (2003), "DyNetML: Interchange Format for Rich Social Network Data", *(Technical Report CMU-ISRI-04-105) CASOS*, Carnegie Mellon University, School of Computer Science, Institute for Software Research International.

[30] i2 Analysts Notebook 7 Product Overview (online), i2 Ltd., http://www.i2.co.uk/Products/Analysts_Notebook (Access Date: 15 January 2008).

[31] i2 Ltd., (October 2007), "NATO Places Follow-on Order for i2 Analyst's Notebook and iBridge" (online), http://www.i2.co.uk/company/press/default.asp?action=view&id=89 (Access Date: 23 January 2008).

[32] Associated Press (December 2003), "Computer-Sleuthing Aids Troops in Iraq" (online), CNN, http://www.cnn.com/2003/TECH/ptech/12/23/tracking.rebels.ap/index.html (Access Date: 23 January 2008).

[33] Plaisant, C., (2004), "The Challenge of Information Visualization Evaluation", In *AVI '04: Proceedings of the Working Conference on Advanced Visual Interfaces*, pp. 109-116, New York, NY, USA: ACM.

[34] Fekete, J.-D. and Plaisant, C., (2006), "Information Visualization Benchmarks Repository" (online), University of Maryland, http://www.cs.umd.edu/hcil/InfovisRepository (Access Date: 12 February 2008).

[35] Lincoln Laboratory, (2001), "DARPA Intrusion Detection Evaluation" (online), Massachusetts Institute of Technology, http://www.ll.mit.edu/IST/ideval (Access Date: 12 February 2008).

[36] Cheswick, B., (2006), "Internet Mapping Project Data" (online), http://cheswick.com/ches/map/dbs (Access Date: 12 February 2008).

[37] Di Battista, G., Garg, A., Liotta, G., Tamassia, R., Tassinari, E. and Vargiu, F., (1995), "An Experimental Comparison of Three Graph Drawing Algorithms", In *Proceedings of the 11th Annual ACM Symposium on Computational Geometry*, pp. 306-315.

[38] Cornell University, (2003), "KDD Cup 2003 Datasets" (online), http://www.cs.cornell.edu/projects/kddcup/datasets.html (Access Date: 12 February 2008).

[39] Chakrabarti, D. and Faloutsos, C., (2006), "Graph Mining: Laws, Generators, and Algorithms", *ACM Computing Surveys*, 38(1).

[40] Shneiderman, B., (1996), "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations", In *VL '96: Proceedings of the 1996 IEEE Symposium on Visual Languages*, p. 336, Washington, DC, USA: IEEE Computer Society.

[41] Amar, R., Eagan, J. and Stasko, J., (2005), "Low-Level Components of Analytic Activity in Information Visualization", In *INFOVIS'05: Proceedings of the 2005 IEEE Symposium on Information Visualization*, p. 15, Washington, DC, USA: IEEE Computer Society.

[42] Amar, R.A. and Stasko, J.T., (2005), “Knowledge Precepts for Design and Evaluation of Information Visualizations”, *IEEE Transactions on Visualization and Computer Graphics*, 11(4), 432-442.

[43] Chi, E.H., (2000), “A Taxonomy of Visualization Techniques Using the Data State Reference Model”, In *INFOVIS ‘00: Proceedings of the IEEE Symposium on Information Visualization 2000*, p. 69, Washington, DC, USA: IEEE Computer Society.

[44] Pattison, T., Vernik, R. and Phillips, M., (2001), “Information Visualisation Using Composable Layouts and Visual Sets”, In *APVis ‘01: Proceedings of the 2001 Asia-Pacific Symposium on Information Visualisation*, pp. 1-10, Darlinghurst, Australia, Australia: Australian Computer Society, Inc.

[45] Schulz, H.-J. and Schumann, H., (2006), “Visualizing Graphs – A Generalized View”, In *IV ‘06: Proceedings of the Conference on Information Visualization*, pp. 166-173, Washington, DC, USA: IEEE Computer Society.

[46] Lee, B., Plaisant, C., Parr, C.S., Fekete, J.-D. and Henry, N., (2006), “Task Taxonomy for Graph Visualization”, In *BELIV ‘06: Proceedings of the 2006 AVI Workshop on Beyond Time and Errors*, pp. 1-5, New York, NY, USA: ACM Press.

[47] Vernik, R., (2000), “A Proposed Reference Model Framework for the Application of Computer-Based Visualisation Approaches”, In *NATO Workshop on Visualisation of Massive Military Datasets*, Québec, Canada.

[48] Kreuseler, M., Lopez, N. and Schumann, H., (2000), “A Scalable Framework for Information Visualization”, In *INFOVIS ‘00: Proceedings of the IEEE Symposium on Information Vizualization 2000*, p. 27, Washington, DC, USA: IEEE Computer Society.

[49] Fujishiro, I., Furuhata, R., Ichikawa, Y. and Takeshima, Y., (2000), “GADGET/IV: A Taxonomic Approach to Semi-Automatic Design of Information Visualization Applications Using Modular Visualization Environment”, In *INFOVIS ‘00: Proceedings of the IEEE Symposium on Information Vizualization 2000*, p. 77, Washington, DC, USA: IEEE Computer Society.

[50] Gouin, D., Vernik, R.J., Evdokiou, P., Houghton, P., Barnes, M., Monk, D. and Rosenblum, L., (2002), “Information Visualization in C3I: Final Report of TTCP C3I AG-3”, *Technical Report DOC-C3I-AG3-1-2002,* TTCP C3I AG-3.

[51] Kreuseler, M. and Schumann, H., (2002), “A Flexible Approach for Visual Data Mining”, *IEEE Transactions on Visualization and Computer Graphics*, 8(1), 39-51.

[52] North, C., Conklin, N. and Saini, V., (2002), “Visualization Schemas for Flexible Information Visualization”, In *INFOVIS ‘02: Proceedings of the IEEE Symposium on Information Visualization*, p. 15, Washington, DC, USA: IEEE Computer Society.

[53] Ross, G. and Chalmers, M., (2003), “A Visual Workspace for Constructing Hybrid Multidimensional Scaling Algorithms and Coordinating Multiple Views”, *Information Visualization*, 2(4), 247-257.

[54] Seo, J. and Shneiderman, B., (2004), “A Rank-by-Feature Framework for Unsupervised Multidimensional Data Exploration Using Low Dimensional Projections”, In *INFOVIS ‘04: Proceedings of the IEEE Symposium on Information Visualization*, pp. 65-72, Washington, DC, USA: IEEE Computer Society.

[55] Tory, M. and Moller, T., (2004), “Rethinking Visualization: A High-Level Taxonomy”, In *INFOVIS ‘04: Proceedings of the IEEE Symposium on Information Visualization*, pp. 151-158, Washington, DC, USA: IEEE Computer Society.

[56] Heer, J., Card, S.K. and Landay, J.A., (2005), “Prefuse: A Toolkit for Interactive Information Visualization”, In *CHI ‘05: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 421-430, New York, NY, USA: ACM Press.

[57] Vernik, R., Itzstein, G.S.V. and Bouchard, A., (2007), “Imago: An Integrated Prototyping, Evaluation and Transitioning Environment for Information Visualisation”, In *11th International Conference on Information Visualisation*, pp. 17-22, Zürich, Switzerland.

[58] Huang, X., Eades, P. and Lai, W., (2005), “A Framework of Filtering, Clustering and Dynamic Layout Graphs for Visualization”, In *ACSC ‘05: Proceedings of the Twenty-Eighth Australasian Conference on Computer Science*, pp. 87-96, Darlinghurst, Australia, Australia: Australian Computer Society, Inc.

[59] Bouchard, A. and Vernik, R., (2006), “Characterisation and Showcasing of Network Visualisation Approaches for Command and Control”, In *Proceedings of Workshop IST-063/RWS-010 Visualising Network Data*, Copenhagen, p. 15, Copenhagen, Denmark, 17-20 October 2006.

[60] Tamassia, R., Battista, G.D. and Batini, C., (1988), “Automatic Graph Drawing and Readability of Diagrams”, *IEEE Trans. Syst. Man Cybern.*, 18(1), 61-79.

[61] Marshall, S., (2007), “Graph Visualization Framework” (online), http://gvf.sourceforge.net Access Date: 18 February 2008).

[62] School of Library and Information Science, Indiana University, (2005), “InfoVis CyberInfrastructure” (online), http://iv.slis.indiana.edu/sw (Access Date: 18 February 2008).

[63] Auber, D., (2008), “Tulip Software” (online), http://tulip-software.org (Access Date: 14 February 2008).

[64] Technical University of Dortmund, University of Cologne, and Oreas GmbH, (2007), “OGDF – Open Graph Drawing Framework” (online), http://www.ogdf.net/ogdf.php (Access Date: 18 February 2008).

[65] The Mathworks, Inc., (2008), “MATLAB and Simulink for Technical Computing” (online), http://www.mathworks.com (Access Date: 14 February 2008).

[66] Wolfram Research Inc., (2008), “Wolfram Mathematica: Home Page” (online), http://www.wolfram.com/products/mathematica (Access Date: 14 February 2008).

[67] Hagberg, A., Schult, D. and Swart, P., (2008), “NetworkX” (online), https://networkx.lanl.gov/wiki (Access Date: 14 February 2008).

[68] Ellson, J., and Gansner, E., (2008), “Graphviz” (online), http://www.graphviz.org (Access Date: 9 July 2008).

[69] Blythe, J., McGrath, C. and Krackhardt, D., (1996), “The Effect of Graph Layout on Inference from Social Network Data”, In Brandenburg, F.J., (Ed.), *Graph Drawing, Passau, Germany, September 20-22, 1995*, pp. 40-51, Springer.

[70] NATO IST-021/RTG-007, (2007), "Multimedia Visualisation of Massive Military Datasets", *Technical Report, RTO-TR-IST-021*, NATO Research and Technology Organisation.

[71] Google Inc., (2007), "Google Earth" (online), http://earth.google.com (Access Date: 24 January 2008).

[72] Sarkar, M. and Brown, M.H., (1994), "Graphical Fisheye Views", *Communications of the ACM*, 37(12), 73-83.

[73] Tominski, C., Abello, J., van Ham, F. and Schumann, H., (2006), "Fisheye Tree Views and Lenses for Graph Visualization", In *IV '06: Proceedings of the Conference on Information Visualization*, pp. 17-24, Washington, DC, USA: IEEE Computer Society.

[74] Lamping, J. and Rao, R., (1994), "Laying Out and Visualizing Large Trees using a Hyperbolic Space", In *UIST '94: Proceedings of the 7th Annual ACM Symposium on User Interface Software and Technology*, pp. 13-14, New York, NY, USA: ACM.

[75] Munzner, T., (1997), "H3: Laying Out Large Directed Graphs in 3D Hyperbolic Space", In *INFOVIS '97: Proceedings of the 1997 IEEE Symposium on Information*, p. 2, Washington, DC, USA: IEEE Computer Society.

[76] AbsInt Angewandte Informatik GmbH, (2007), "aiSee Graph Layout Software" (online), http://www.aisee.com (Access Date: 19 February 2008).

[77] Munzner, T., (2001), "H3Viewer" (online), http://graphics.stanford.edu/~munzner/h3 (Access Date: 19 February 2008).

[78] Fekete, J.-D., (2005), "The InfoVis Toolkit" (online), http://ivtk.sourceforge.net (Access Date: 19 February 2008).

[79] Baudisch, P., Good, N., Bellotti, V. and Schraedley, P., (2002), "Keeping Things in Context: A Comparative Evaluation of Focus plus Context Screens, Overviews, and Zooming", In *CHI '02: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 259-266, New York, NY, USA: ACM.

[80] Gutwin, C. and Fedak, C., (2004), "Interacting with Big Interfaces on Small Screens: A Comparison of Fisheye, Zoom, and Panning Techniques", In *GI '04: Proceedings of Graphics Interface 2004*, pp. 145-152, School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada: Canadian Human-Computer Communications Society.

[81] Hollands, J.G., Carey, T.T., Matthews, M.L. and McCann, C.A., (1989), "Presenting a Graphical Network: A Comparison of Performance using Fisheye and Scrolling Views", In *Proceedings of the Third International Conference on Human-Computer Interaction on Designing and using Human-Computer Interfaces and Knowledge Based Systems (2nd ed.)*, pp. 313-320, New York, NY, USA: Elsevier Science Inc.

[82] Gansner, E.R. and Koren, Y., (2007), "Improved Circular Layouts", In Kaufmann, M. and Wagner, D., (Eds.), *Graph Drawing, Karlsruhe, Germany, September 18-20, 2006*, pp. 386-398, Springer.

[83] Holten, D., (2006), "Hierarchical Edge Bundles: Visualization of Adjacency Relations in Hierarchical Data", *IEEE Transactions on Visualization and Computer Graphics*, 12(5), 741-748.

[84] Dill, J., Bartram, L., Ho, A. and Henigmann, F., (1994), "A Continuously Variable Zoom for Navigating Large Hierarchical Networks", In *Proceedings of the IEEE Conference on Systems, Man and Cybernetics*, pp. 386-390.

[85] Jgraph, (2007), "JGraph Features" (online), http://www.jgraph.com/jgraphfeatures.html (Access Date: 28 January 2008).

[86] Wu, A.Y., Garland, M. and Han, J., (2004), "Mining Scale-Free Networks using Geodesic Clustering", In *KDD '04: Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 719-724, New York, NY, USA: ACM.

[87] Bjørke, J.T., (2006), "Reduction of Complexity: An Aspect of Network Visualization", In *Proceedings of IST-063/RWS-010 Visualising Network Information*, Copenhagen, DK.

[88] Huang, W., Murray, C., Shen, X., Song, L., Wu, Y.X. and Zheng, L., (2005), "Visualisation and Analysis of Network Motifs", In *IV '05: Proceedings of the Ninth International Conference on Information Visualisation*, pp. 697-702, Washington, DC, USA: IEEE Computer Society.

[89] Klukas, C., Schreiber, F. and Schwöbbermeyer, H., (2006), "Coordinated Perspectives and Enhanced Force-Directed Layout for the Analysis of Network Motifs", In *APVIS '06: Proceedings of the Asia Pacific Symposium on Information Visualisation*, pp. 39-48, Darlinghurst, Australia, Australia: Australian Computer Society, Inc.

[90] Moody, D., (2007), "What Makes a Good Diagram? Improving the Cognitive Effectiveness of Diagrams in IS Development", In *Advances in Information Systems Development: New Methods and Practice for the Networked Society*, Vol. 2, pp. 481-492, Springer.

[91] Abello, J. and Korn, J., (2000), "Visualizing Massive Multi-Digraphs", In *INFOVIS '00: Proceedings of the IEEE Symposium on Information Vizualization 2000*, p. 39, Washington, DC, USA: IEEE Computer Society.

[92] Teoh, S.T. and Ma, K.-L., (2002), "RINGS: A Technique for Visualizing Large Hierarchies", In *GD '02: Revised Papers from the 10th International Symposium on Graph Drawing*, pp. 268-275, London, UK: Springer-Verlag.

[93] Stravers, P., (1991), "Partitioning a Network into n Pieces with a Time-Efficient Net Cost Function", In *EURO-DAC '91: Proceedings of the Conference on European Design Automation*, pp. 177-182, Los Alamitos, CA, USA: IEEE Computer Society Press.

[94] Auber, D., Chiricota, Y., Jourdan, F. and Melançon, G., (2003), "Multiscale Visualization of Small World Networks", In *INFOVIS'03: Proceedings of the IEEE Symposium on Information Visualization*, pp. 75-81.

[95] Dwyer, T. and Koren, Y., (2005), "DIG-COLA: Directed Graph Layout through Constrained Energy Minimization", In *INFOVIS '05: Proceedings of the Proceedings of the 2005 IEEE Symposium on Information Visualization*, p. 9, Washington, DC, USA: IEEE Computer Society.

[96] Vandenberghe, G., (2006), "Automating the Presentation of Computer Networks", In *Proceedings of IST-063/RWS-010 Visualising Network Information*, RTO-MP IST-063, Copenhagen, Denmark.

[97] Walshaw, C., (2003), “A Multilevel Algorithm for Force-Directed Graph-Drawing”, *Journal of Graph Algorithms and Applications*, 7(3), 253-285.

[98] Chan, D., Chua, K., Leckie, C. and Parhar, A., (2003), “Visualisation of Power-Law Network Topologies”, In *Proceedings of the Eleventh IEEE International Conference on Networks (ICON 2003)*, pp. 69-74, Sydney, Australia.

[99] Ghoniem, M., Fekete, J.-D. and Castagliola, P., (2004), “A Comparison of the Readability of Graphs Using Node-Link and Matrix-Based Representations”, In *INFOVIS'04: Proceedings of the IEEE Symposium on Information Visualization*, pp. 17-24, Washington, DC, USA: IEEE Computer Society.

[100] Henry, N. and Fekete, J.-D., (2006), “MatrixExplorer: A Dual-Representation System to Explore Social Networks”, *IEEE Transactions on Visualization and Computer Graphics*, 12(5), 677-684.

[101] Howes, A. and de Bruijn, O., (2005), “A Review of Approaches to Displaying Uncertain and Complex Information”, Technical Report.

[102] Houston, B., Lem, M., Taylor, V., Kaster, A., Treurniet, J. and Jacobson, Z., (2005), “Representing Uncertainty, Unknowns and Dynamics of Social Networks”, In *N/X-Workshop on Social Network Analysis and Visualization for Public Safety*, FGAN-FKIE, Wachtberg-Werthhoven, Germany.

[103] Bilgic, M., Licamele, L., Getoor, L. and Shneiderman, B., (2005), “D-Dupe: An Interactive Tool for Entity Resolution in Social Networks”, In Healy, P. and Nikolov, N.S., (Eds.), *International Symposium on Graph Drawing*, Vol. 3843 of *Lecture Notes in Computer Science*, pp. 505-507, Springer.

[104] TTCP C3I TP2, (2006), “Large Group Displays in Command and Control Environments: Knowledge Integration Task (KnowIT) Results”, *Technical Report TR-C3I-TP2-2-2006*.

[105] DynaScan, (2007), “360 Degree LED Video Displays” (online), DynaScan Technology Corp., http://www.dynascan.com.tw (Access Date: 3 October 2007).

[106] Microsoft, (2007), “Microsoft Surface” (online), Microsoft Corp., http://www.microsoft.com/surface/ (Access Date: 3 October 2007).

[107] Frishman, Y. and Tal, A., (2007), “Multi-Level Graph Layout on the GPU”, *IEEE Transactions on Visualization and Computer Graphics*, 13(6), 1310-1319.

[108] NetSci International Workshop and Conference on Network Science, (2007), “Competition on Visualizing Network Dynamics” (online), http://vw.indiana.edu/07netsci/entries (Access Date: 19 February 2008).

[109] Kapler, T. and Wright, W., (2005), “Geo Time Information Visualization”, *Information Visualization*, 4(2), 136-146.

[110] ILOG Inc., (2008), “ILOG JViews Diagrammer: Java and Ajax Diagrams and Dashboards” (online), http://www.ilog.com/products/jviews/diagrammer (Access Date: 19 February 2008).

[111] Nevron, (2007), “Nevron Diagram for .NET” (online), http://www.nevron.com/Products.DiagramFor.NET.Overview.aspx (Access Date: 19 February 2008).

[112] yWorks, (2008), "yFiles – Java Graph Layout and Visualization Library" (online), http://www.yworks.com/en/products_yfiles_about.html (Access Date: 19 February 2008).

[113] Tom Sawyer Software, "Tom Sawyer Software" (online), http://www.tomsawyer.com (Access Date: 15 January 2008).

[114] Moody, J., McFarland, D.A. and Bender-deMoll, S., (2005), "Visualizing Network Dynamics", *American Journal of Sociology*, 110(4), 1206-1241.

[115] Brandes, U. and Corman, S.R., (2002), "Visual Unrolling of Network Evolution and the Analysis of Dynamic Discourse", In *INFOVIS '02: Proceedings of the IEEE Symposium on Information Visualization*, p. 145, Washington, DC, USA: IEEE Computer Society.

[116] Nakazono, N., Misue, K. and Tanaka, J., (2006), "NeL2: Network Drawing Tool for Handling Layered Structured Network Diagram", In *APVis '06: Proceedings of the 2006 Asia-Pacific Symposium on Information Visualisation*, pp. 109-115, Darlinghurst, Australia, Australia: Australian Computer Society, Inc.

[117] Frishman, Y. and Tal, A., (2004), "Dynamic Drawing of Clustered Graphs", In *INFOVIS '04: Proceedings of the IEEE Symposium on Information Visualization*, pp. 191-198, Washington, DC, USA: IEEE Computer Society.

[118] Kumar, G. and Garland, M., (2006), "Visual Exploration of Complex Time-Varying Graphs", *IEEE Transactions on Visualization and Computer Graphics*, 12(5), 805-812.

[119] Jucknath-John, S. and Graf, D., (2006), "Icon Graphs: Visualizing the Evolution of Large Class Models, In *SoftVis '06: Proceedings of the 2006 ACM Symposium on Software Visualization*, pp. 167-168, New York, NY, USA: ACM.

[120] University of Arizona, "GraphAEL" (online), http://graphael.cs.arizona.edu/graphael (Access Date: 15 January 2008).

[121] Stanford University, "Social Network Image Animator" (online), http://www.stanford.edu/group/sonia (Access Date: 15 January 2008).

[122] MIT, "iCKN – Innovative Collaborative Knowledge Networks" (online), http://www.ickn.org (Access Date: 15 January 2008).

[123] Chen, C. and Morris, S., (2003), "Visualizing Evolving Networks: Minimum Spanning Trees versus Pathfinder Networks", In *INFOVIS '03: Proceedings of the IEEE Symposium on Information Visualization*, pp. 67-74, Washington, DC, USA: IEEE Computer Society.

[124] Börner, K., Sanyal, S. and Vespignani, A., (2007), "Network Science", *Annual Review of Information Science and Technology*, 41, 537-607.

[125] Pigeon, L., Lecornu, L., Bergeron-Guyard, A. and Bergeron, A., (2005), "Navigation in Extreme Environments: SCIPIO Optipath", *DRDC Valcartier TR-2005-529*, Defence R&D Canada – Valcartier.

[126] Perer, A. and Shneiderman, B., (2008), "Integrating Statistics and Visualization: Case Studies of Gaining Clarity during Exploratory Data Analysis", To appear in the *Proceedings of CHI2008*, Florence, Italy.

[127] Sawilla, R. and Ou, X., (2007), “Googling Attack Graphs”, *DRDC Ottawa TM-2007-205*, Defence R&D Canada – Ottawa.

[128] Kules, B., Shneiderman, B. and Plaisant, C., (2003), “Data Exploration with Paired Hierarchical Visualizations: Initial Designs of Pair Trees”, In *DG.O’03: Proceedings of the 2003 Annual National Conference on Digital Government Research*, Digital Government Research Center.

[129] Andrienko, N. and Andrienko, G., (2003), “Informed Spatial Decisions through Coordinated Views”, *Information Visualization*, 2, 270-285.

[130] Namata, G.M., Staats, B., Getoor, L. and Shneiderman, B., (2007), “A Dual-View Approach to Interactive Network Visualization”, In *CIKM ‘07: Proceedings of the Sixteenth ACM Conference on Information and Knowledge Management*, pp. 939-942, New York, NY, USA: ACM.

[131] Andrienko, G. and Andrienko, N., (2007), “Coordinated Multiple Views: A Critical View”, In *CMV ‘07: Proceedings of the Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization*, pp. 72-74, Washington, DC, USA: IEEE Computer Society.

[132] Taylor, R., (2002), “Visualizing Multiple Fields on the Same Surface”, *IEEE Computer Graphics and Applications*, 22(3), 6-10.

[133] Treurniet, J., (2005), “Social Cyber Networks”, *N/X-Workshop on Social Network Analysis and Visualization for Public Safety*, FGAN-FKIE, Wachtberg-Werthhoven, Germany.

[134] Thaler, D. and Ravishankar, C.V., (1995), “NView: A Visual Framework for Network Tool Integration”, In *Proceedings of the 1995 IEEE Fourteenth Annual International Phoenix Conference on Computers and Communications*, pp. 283-289, Phoenix, AZ, USA: IEEE Computer Society.

[135] Shneiderman, B. and Aris, A., (2006), “Network Visualization by Semantic Substrates”, *IEEE Transactions on Visualization and Computer Graphics*, 12(5), 733-740.

[136] Erten, C., Kobourov, S.G., Le, V. and Navabi, A., (2004), “Simultaneous Graph Drawing: Layout Algorithms and Visualization Schemes”, In Liotta, G., (Ed.), *Graph Drawing, Perugia, 2003*, pp. 437-449, Springer.

[137] Williams, M. and Munzner, T., (2004), “Steerable, Progressive Multidimensional Scaling”, In *INFOVIS ‘04: Proceedings of the IEEE Symposium on Information Visualization*, pp. 57-64, Washington, DC, USA: IEEE Computer Society.

[138] Purchase, H.C., Cohen, R.F. and James, M.I., (1997), “An Experimental Study of the Basis for Graph Drawing Algorithms”, *Journal of Experimental Algorithmics*, 2, 4.

[139] Xu, J.J. and Chen, H., (2005), “CrimeNet Explorer: A Framework for Criminal Network Knowledge Discovery”, *ACM Transactions on Information Systems*, 23(2), 201-226.

[140] North, C., (2006), “Toward Measuring Visualization Insight”, *IEEE Computer Graphics and Applications*, 26(3), 6-9.

[141] Cawthon, N. and Vande Moere, A., (2007), "The Effect of Aesthetic on the Usability of Data Visualization", In *11th International Conference on Information Visualisation*, pp. 637-648, Zürich, Switzerland.

[142] Huang, W., Hong, S.-H. and Eades, P., (2007), "Effects of Sociogram Drawing Conventions and Edge Crossings in Social Network Visualization", *Journal of Graph Algorithms and Applications*, 11(2), 397-429.

[143] Python Software Foundation, (2008), "Python Programming Language" (online), http://www.python.org (Access Date: 14 February 2008).

[144] iDashboards, (2007), "Business Intelligence Dashboard Software, Metrics, Scorecard, KPI, Alert, iDashboard" (online), http://idashboards.com (Access Date: 14 February 2008).

[145] McGookin, D.K. and Brewster, S.A., (2006), "SoundBar: Exploiting Multiple Views in Multimodal Graph Browsing", In *NordiCHI '06: Proceedings of the 4th Nordic Conference on Human-Computer Interaction*, pp. 145-154, New York, NY, USA: ACM.

[146] Papadopoulos, C., Kyriakakis, C., Sawchuk, A. and He, X., (2004), "CyberSeer: 3D Audio-Visual Immersion for Network Security and Management", In *VizSEC/DMSEC '04: Proceedings of the 2004 ACM Workshop on Visualization and Data Mining for Computer Security*, pp. 90-98, New York, NY, USA: ACM.

[147] Harrop, W. and Armitage, G., (2006), "Real-Time Collaborative Network Monitoring and Control using 3D Game Engines for Representation and Interaction", In *VizSEC '06: Proceedings of the 3rd International Workshop on Visualization for Computer Security*, pp. 31-40, New York, NY, USA: ACM.

[148] .theprodukkt GmbH, (2007), ".werkkzeug 3 TE" (online), http://www.werkkzeug.com (Access Date: 14 February 2008).

[149] Pacific Northwest National Laboratory, (2007), "Starlight Information Visualization Technologies" (online), http://starlight.pnl.gov (Access Date: 22 February 2008).

[150] Stasko, J., Goerg, C., Liu, Z., Stepanov, G., Panatazopoulos, V., Williams, S., Parekh, N. and Singhal, K., (2007), "Jigsaw: Visualization for Investigative Analysis" (online), Georgia Tech, http://www.gvu.gatech.edu/ii/jigsaw (Access Date: 22 February 2008).

[151] Rhyne, T.-M., (2000), "Computer Games' Influence on Scientific and Information Visualization", *IEEE Computer*, 33(12), 154-156.

[152] Matthews, M.L., Webb, R.D.G. and McCann, C., (1997), "A Framework for Evaluation of Military Command and Control Systems", *DCIEM-97-CR-24*, Defence Civil Institute of Environmental Medicine.

Chapter 4 – HYPERNODES: THEIR APPLICATION TO VISUALISATION OF NETWORKS

4.1 INTRODUCTION

Graphs are often used in the modelling of real world network-related phenomena, such as road networks, computer networks, social networks, the representation of abstract concepts or in the analysis of systems, documents, images, etc. The reduction of complexity is an important element in the management and visualisation of networks with large numbers of nodes and links. Bjørke [6] shows how this idea leads to the concept of hypernodes. Networks of hypernodes enable networks to be represented at different levels of abstraction, as in an interactive map that allows the user to zoom in and out so that the level of detail is adapted to the desired map scale. In other words it transforms a flat network into a hierarchical struture that can reveal or highlight the underlying structure/pattern of the network in an effective and intuitive manner.

Bertin [3] introduced a graphical method to find groups in geographical data. Initially, the data is mapped to an image, which Bertin termed the reorderable matrix. Then the rows, or columns, of this image are interchanged to generate different views of the data. In this way meaningful patterns in the data can be detected by visual interpretation of the reordered image. Bertin also constructed a mechanical permutation technique ([3], p. 35). Based on the idea of the reorderable matrix Siirtola and Mäkinen [16] present a tool for interactive cluster analysis. Bjørke and Smith [5] developed an algorithm to automate the reorganization (also termed seriation) of the reorderable matrix in which the seriation criterion is defined on the basis of the minimum entropy of a binary image. Based on reorganization of the adjacency matrix of networks, an automated method to construct hierarchies of networks can be formulated [6].

IST-059 has identified key issues for the visualisation of networks and points out that fuzziness and uncertainty are aspects that must be considered in real-world network visualisation. Indeed, uncertainty is unavoidable in networks and this posses a challenge in analysing and visualising the network. As the number of the nodes and links increases, compounded with uncertainty, the representation of the network needs simplification in order to keep the visual clarity of the image while taking into account, for example, propagating, the degree of the uncertainties and their effects on the topological structure of the network [17]. There is, therefore, a need to extend the certainty-based hypernode algorithm to handle uncertain relationships [7].

There are two main categories of uncertainties in a network, namely uncertainties about the edges and uncetainties about the nodes; both will be discussed in this chapter.

If uncertainty about an edge could be mapped to a membership function in a class such as "perfect edge", the concept of fuzzy relations could be applied. Crisp relations can be described by their characteristic function, i.e. an edge in a crisp network is associated with the number 1 or 0 dependent on whether the edge exists or not. In a fuzzy relation (binary) the edges are allowed to have varying degrees of membership within the relation (see for example [13], page 120). Although uncertainty is quite distinct from fuzziness, such a mapping may be admissible in many cases.

The terms hyperedge and hypegraph are used in mathematical literature. In order not to confuse with the established theory of hypergraphs [[1], [2], [18] and [19]], we will not use these terms. Huang and Lai [11] cluster nodes in graphs and apply a method which is parallel to our hypernode concept. They use the terms abstract node, supernode and metanode. For edges among abstract nodes they use the term abstract edge. Flake et al. [10] also demonstrate a similar concept. A comprehensive introduction to network science can be found, for example, in Borner et al. [8].

In this chapter the extension of the crisp based hypernode algorithm to the case of weighted graphs is described and discussed. “Weight” may apply eaither to uncertainty or to fuzzy membership (Annex B, Section B.2.3.1). Some examples are given to demonstrate how the hypernode algorithm maps nodes and links to networks of hypernodes.

4.2 METHOD TO CONSTRUCT NETWORKS OF HYPERNODES

4.2.1 Crisp Networks

Figure 4-1 shows an adjacency matrix that represents a crisp network. If there is a connection between two nodes in the network, the corresponding cell in the image of the Adjacency Matrix is coloured (in our case blue), else it is white. Numerically, the binary property of the matrix can be represented by the numbers 1 or 0, i.e. a characteristic function.

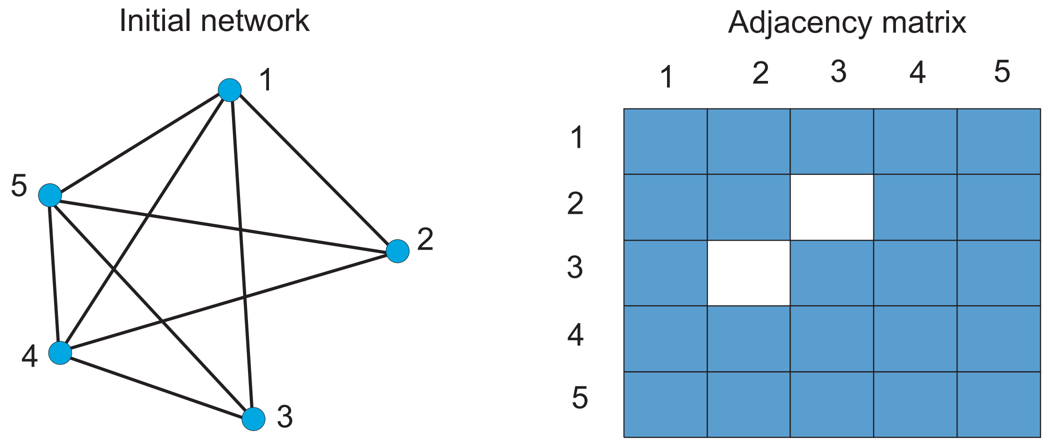


Figure 4-1: A Network and its Adjacency Matrix. If there is a connection between two nodes in the network, the corresponding cell in the adjacency matrix is coloured blue (i.e. a value of 1), else it is white.

The adjacency matrix of a network can be mapped to an image so that similar rows, or similar columns, are clustered. Figure 4-1 and Figure 4-2 demonstrate how reordering can be used to get a view of the adjacency matrix where groups of nodes can be derived. From the reordered matrix, three groups of nodes can be derived, i.e. the hypernodes H1, H2 and H3 as shown in Figure 4-2. Hypernode H1 represents an aggregation of the tree sub-nodes 1, 4 and 5. From the network of hypernodes we can see that the initial network in Figure 4-1 can be regarded as a tree structure, one mother node and two leaf nodes. The mother node H1 is composed of three sub-nodes (1, 4 and 5) which have strong connection, i.e. one node is connected to the other two. Since there is no direct connection between nodes 2 and 3, hypernodes H2 and H3 have no direct link. The network in Figure 4-2 is constructed by applying the forthcoming algorithms (see below) as shown in Figure 4-3 and Figure 4-4. The size of the node is proportional to its connectivity. Indeed Hypernode H1 can be interpreted as the key influential node in the graph while Hypernodes H2 and H3 represent the sub-graphs in a hierechical relationship in this representation.

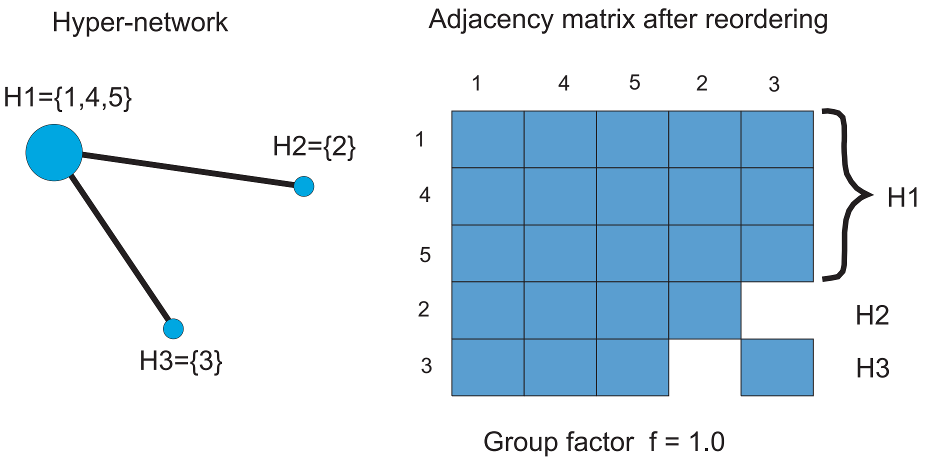


Figure 4-2: The Adjacency Matrix in Figure 4-1 after Reordering. The group factor f is set to 1, i.e. rows must have similarity 1 in order to be aggregated to a hypernode. The rows are labeled 1,4,5,2 and 3. Row 2, for example, refers to row labeled 2.

```
k = max(sum(R.a(1:n,:),2)); % find the maximum row
R = swap_rows(R,1,k); % move the maximum row to the top of the matrix
% move similar rows close to each other
(1)     for i = 1:n-1
        smin = inf; % a large number
(2)     for ii = i+1:n
            % the mother row is row i, the row to be investigated is row ii
(3)     s = sum(abs(R.a(i,:)-R.a(ii,:)));
            if s<smin
                kk = ii; % candidate row
                smin = s;
            end
        end
            % move the candidate row close to the mother row
            R = swap_rows(R,i+1,kk);
        end
```

Figure 4-3: Pseudocode (MATLAB Code) to Reorder the Adjacency Matrix.

```
C = R.a(k,:); % k is the first row of the group
 i = k+1;
 while i <= n
(1)     Q = abs(C-R.a(i,:)); % distance to the first row of the group
        % find the columns where the one or the other row has membership value greater the 0,
        % i.e. the support of the union of the two rows considered.
(2)     U = find(R.a(i,:)>0 | C>0);
(3)     d = sum(Q)/length(U); % normalized distance
(4)     mu = abs(1-d); % how similar the rows are
(5)     if mu<f % strength of group membership
            return
(6)     else
            i = i+1;
        end
        end
```

Figure 4-4: Pseudocode (MATLAB Code) to Define Groups of Rows in the Reordered Adjacency Matrix.

In order to find the global best ordering of the rows or columns, all combinations of rows or columns should be investigated and global similarity measures introduced. The algorithm described in Figure 4-3 gives an approximation to the global best order, since the algorithm assumes that what is best locally is also best globally. For the purpose of the present purposes and demonstration, the Greedy algorithm proposed is assumed to be sufficient.

The time complexity of the algorithm in Figure 4-3 is $T = O(mn^2)$, where m is the number of columns and n is the number of rows. The computing in step (3) runs over all the columns, i.e. this step takes $O(n)$ time. Step (3) is enveloped in the two nested loops of steps (1) and (2). Since the adjacency matrix of a network is a $n \times n$ matrix, the reordering of the rows and the columns takes $T = O(n^3)$ time. If the algorithm is to be applied to huge networks, for example if $n >> 1000$, the computing time should be considered, i.e. implement methods to limit the exponential growth of the computing time. However, this is outside the scope of the present chapter.

Assume an adjacency matrix R of size $n \times n$ with the elements $r_{i,j}$, where i and j represent the row and columns, respectively, and can be any integer in the interval $[1, n]$. If there is a connection from node i to node j, $r_{i,j} = 1$, else $r_{i,j} = 0$. If the graph is undirected, $r_{i,j} = r_{j,i}$.

The distance between two rows i and k can be defined as:

$$d(i,k) = \sum_{j=1}^{n} | r_{i,j} - r_{k,j} | . \tag{1}$$

Two rows are of maximum similarity when $d = 0$ and minimum similarity when $d = n$. For the adjacency matrix in Figure 4-1, we have $d(1,2) = 1$, $d(2,3) = 2$, $d(3,4) = 1$ and $d(4,5) = 0$, i.e. crisp.

The reordering of R based on the distance measure $d(\cdot)$ as defined in Equation 1, can be derived from an information theoretic point of view [4]. The entropy of a binary image can be computed on the basis of the probability that neighbouring pixels have the same colour (p_+) or different colours (p_-). Since the entropy of R gets its minium value when $p_+ = 1$ and $p_- = 0$, reordering R so that $d(\cdot)$ is minimized corresponds to minimizing the entropy of R. The alternative case when $p_+ = 0$ and $p_- = 1$, corresponds to the chess board layout of R. Since the goal of the hypernode algorithm is to cluster nodes which have strong ssimilarities, the alternative case does not represent a solution to the hypernode problem. A question is whether there exists a broad class of similarity measures that can be used to reorganize the matrix. The answer is application dependent and although there types of similarity measures other than the one described in Equation 1, further discussion is outside the scope of the present chapter.

Readers should note that in this chapter “strength” means “fuzzy membership” unless otherwise specified to mean something else, such as “traffic capacity”. The algorithm is, however, agnostic to the meaning of the matrix entries. Annex E suggests potential application of the same algorithm when the matrix entries are arbitrary attributes of the nodes.

When R is reordered, the question of how to define groups of rows in R arises. A pseudocode of a grouping algorithm is shown in Figure 4-4. Here, the distance measure defined in Equation 1 is applied, but it is normalized and inverted, i.e. $d(i,k)$ is transformed to a number in the interval $[0,1]$ so that 1 means maximum similarity and 0 minimum similarity. The similarity $\mu(k,i)$ between two rows (or columns) k and i is defined as:

$$\mu(k,i) = 1 - \frac{d(k,i)}{e}, \tag{2}$$

where e is the number of links in the union set of the two rows (or columns). Equation 2 is implemented in steps (3) and (4) in Figure 4-4.

For example, the similarity between rows 1 and 2 in Figure 4-1 is:

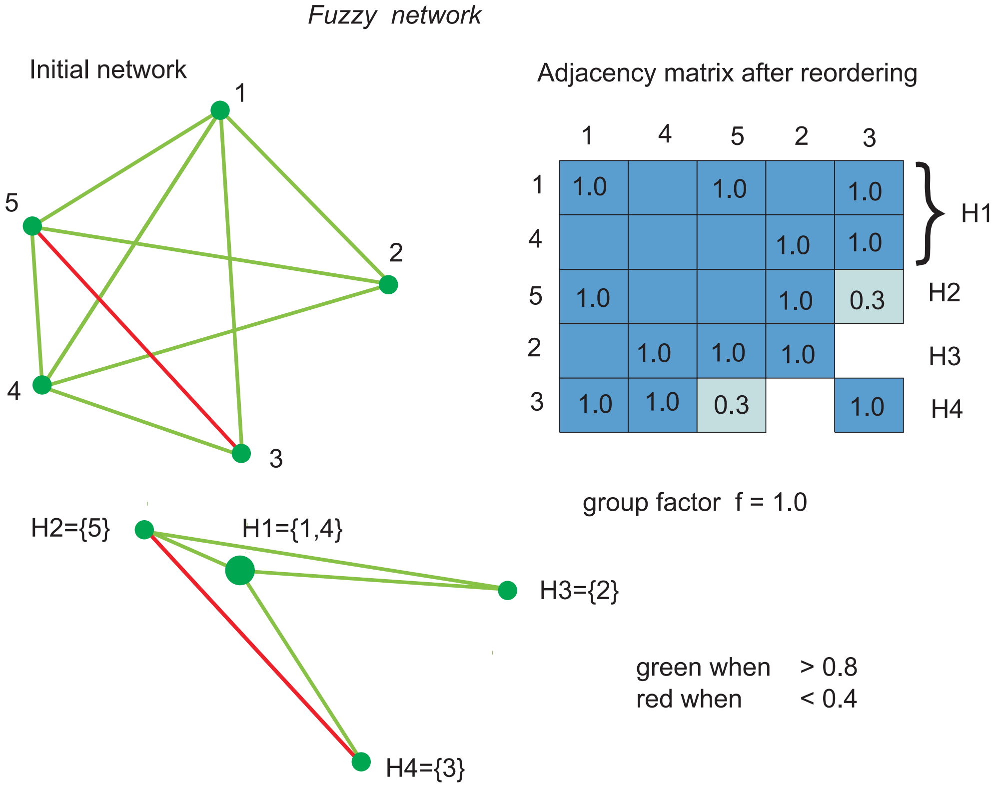
$$\mu(1,2) = 1 - \frac{1}{5} = 1 - 0.2 = 0.8.$$

The computed similarity is compared to a threshold, as shown in steps 5 and 6 of the pseudocode in Figure 4-4. In the example in Figure 4-2 the group factor $f = 1.0$ is applied, i.e. the threshold of the similarity factor is 1.0. Therefore, rows 1 and 2 are not grouped. Their similarity is 0.8, which is less than $f = 1.0$. Row 1 is in this case the first row of the group. The similarity is computed to the first row of the group to be generated, see step (1) in the pseudocode.

4.2.2 Networks with Weighted Edges and Nodes

In the above section we described the application of the hypernode algorithm for a crsip network. In this section we introduce a simple approach to advance the crisp hypernode algorithm to a weighted networks such as a.fuzzy network. This is achieved by replacing the characteristic function with a membership function or a weight function.

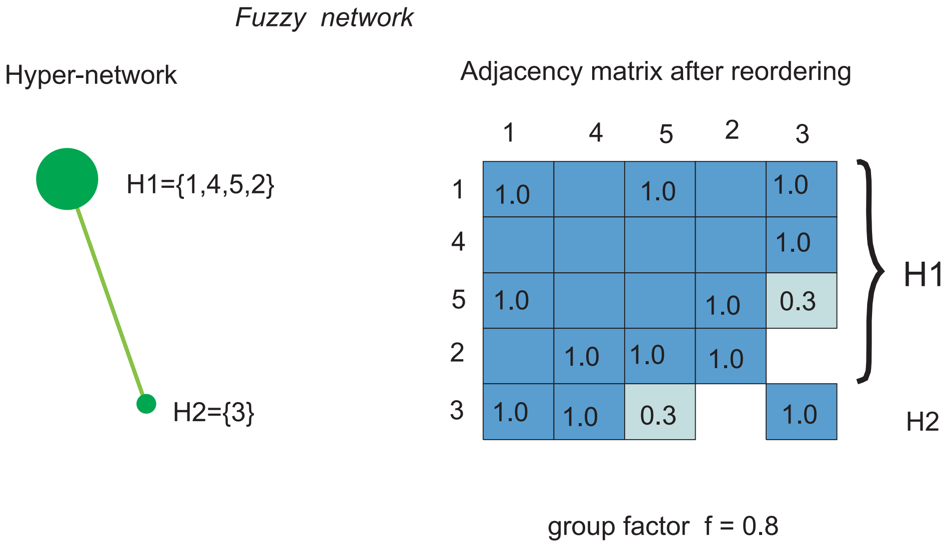
The hypernode algorithm is applied to the reordering and grouping of the adjacency matrix in Figure 4-5. The group factor is set to 1. Figure 4-5 illustrates how fuzziness in a network can be visualised by the application of colour coding. A traffic light colour scheme is used. The green nodes and links have membership value $\mu(r_{i,j}) = 1$ and the red link $\mu(r_{i,j}) = 0.3$. Compared with the crisp network in Figure 4-2, hypernode $H1$ in the weighted (fuzzy) case (Figure 4-5) represents one sub-node less than in the crisp case, i.e. only nodes 1 and 4 are included and not node 5. This is due to the weak link between node 5 and 3. Therefore, node 5 is not aggregated to hypernode $H1$ in the weighted (fuzzy) case when the group factor is set to 1.



	1	4	5	2	3	
1	1.0		1.0		1.0	H1
4				1.0	1.0	
5	1.0			1.0	0.3	H2
2		1.0	1.0	1.0		H3
3	1.0	1.0	0.3		1.0	H4

Figure 4-5: A Weighted Network, its Adjacency Matrix and its Network of Hypernodes. The group factor f is 1.0. The green nodes and links have membership value $\mu(r_{i,j}) = 1$ and the red link $\mu(r_{i,j}) = 0.3$.

The role of H1 changes here due to the fuzzy link between Nodes 3 and 5, i.e. it is no longer the most connected node. The topological structure and relationship are depicted very clearly in this hypernode representation; indeed H2 is the most connected hypernode in this instance. Figure 4-6 illustrates how fuzziness of a network can be visualised by the application of colour coding.



	1	4	5	2	3
1	1.0		1.0		1.0
4					1.0
5	1.0			1.0	0.3
2		1.0	1.0	1.0	
3	1.0	1.0	0.3		1.0

Figure 4-6: The Weighted (Fuzzy) Network in Figure 4-5 Constructed by Group Factor $f = 0.8$.

The group factor determines the structure of the hypernode and must be set with careful consideration in respect to the application in question. Figure 4-6 demonstrates how the group factor, in this example reduced from 1 to 0.8, radically changes the hypernode structure. In this case the network is abstracted to two hypernodes: $H1$ and $H2$.

The above examples show that both the fuzziness and the group factor play an important role in determining the resultant hypernode structure.

4.2.2.1 Visualisation and Representation of Prohibited/Unlikely Links

In the previous section when there is no link between two nodes, for example, nodes 2 and 3 in Figure 4-1, a value 0 is assigned and the corresponding cell is coloured white in the matrix. There is no prior knoweldge as to why there is not a link between them [20]. Indeed the assumption is that there is nothing at all to preclude there being a link between them at all.

However, there are cases where some links are prohibited or highly unlikely. The question is how to differentiate between links that just do not exist on the one hand and links that are prohibited on the other hand, and also how to work with this information. There are many different ways that this can be addressed; one way is through the use of prior beliefs (e.g. that a link cannot exist or is highly unlikely) alongside measurements, for example, as in Bayes' theorem, more details can be found in Annex E. In the example network shown in Figure 4-7 there are no links between nodes 2 and 3 and nodes 2 and 5. Let us assume that there is *a priori* knowledge that it is impossible or prohibited (or at least highly unlikely) to connect nodes 2 and 3.

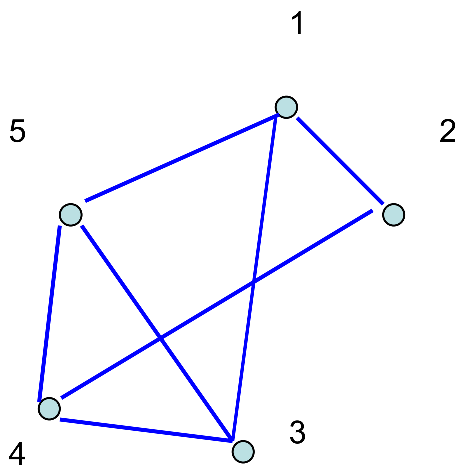


Figure 4-7: Network with No Link between Nodes 2 and 3 and Nodes 2 and 5.

Let us define that:

x = 1 means there is a link

x = 0 means the link is prohibited or highly unlikely

and

y = 1 means a link is observed

y = 0 means a link is not observed

If we assign some reasonable prior probability measures as follows:

$p(x=1) = 0.9$

$p(x=0) = 0.1$

$p(y=0|x=1) = 0.1$

$p(y=1|x=1) = 0.9$

$p(y=0|x=0) = 0.95$

$p(y=1|x=0) = 0.05$

Then applying Bayes' Theorem $P(x|y) = P(y|x)P(x)/\Sigma_x p(y|x)p(x)$ to create posterior probability values we get:

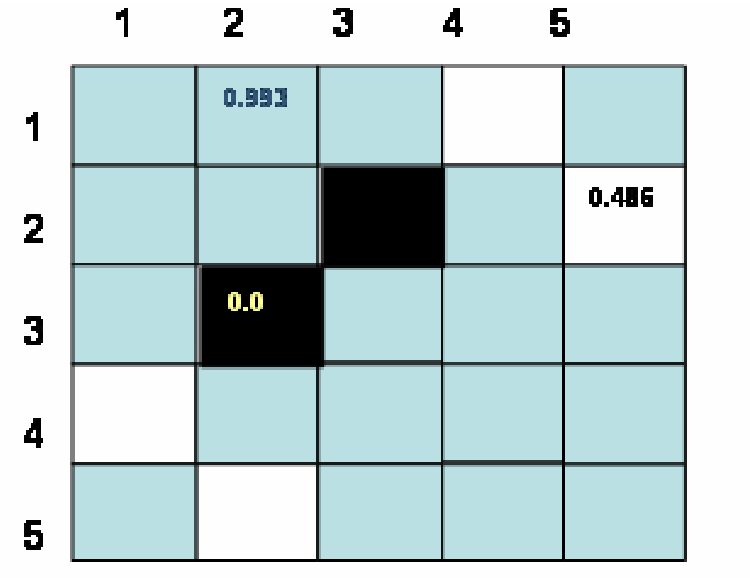
$P(x=1|y=1) = 0.993$

$p(x=1|y=0) = 0.486$

$p(x=0|y=0) = 0.513$

$p(x=0|y=1) = 0.006$

In addition, when a link is highly unlikely the square can be coloured black (nodes 2 and 3) and white when a link (nodes 2 and 5) may exist but is not observed.



	1	2	3	4	5
1		0.993			
2					0.486
3		0.0			
4					
5					

Figure 4-8: Adjacency Matrix with Prohibited Links.

In this way we can still apply the hypernode algorithm but using different measures. Furthermore, we can also still visualise prohibited links within the network (Figure 4-9) and potential missing links within the network (Figure 4-10). Figure 4-11 illustrates how the overall position can be visualised.

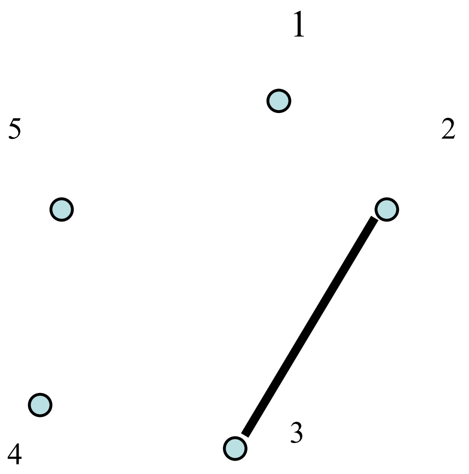


Figure 4-9: A Visualisation of the Prohibited Link.

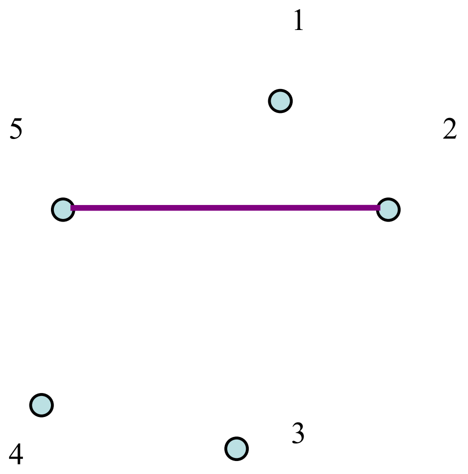


Figure 4-10: A Visualisation of the Potential Missing Link.

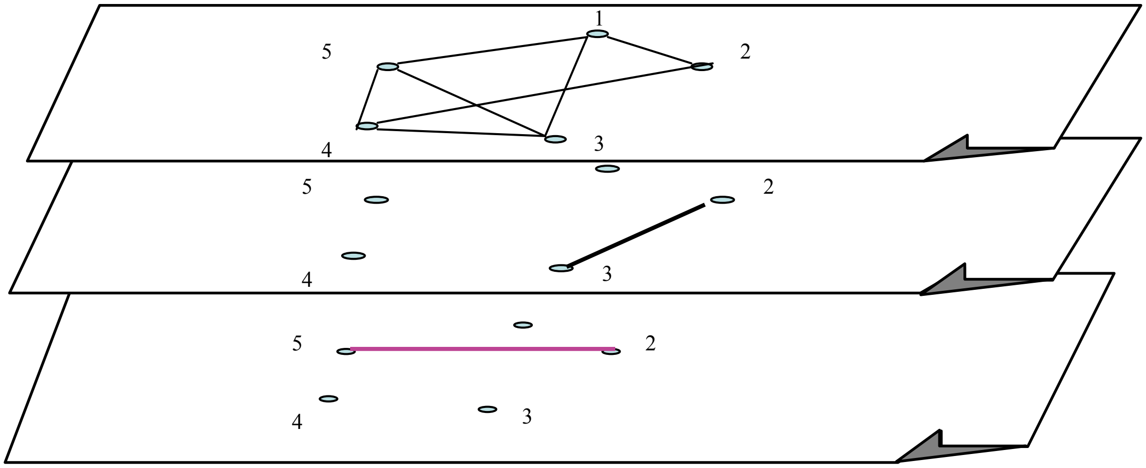


Figure 4-11: Complete Layered Network.

A layered network representation can be visualised by stacking the networks together to give additional information regarding the underlying structure and pattern of the whole network.

4.2.3 Propagation of Uncertainties

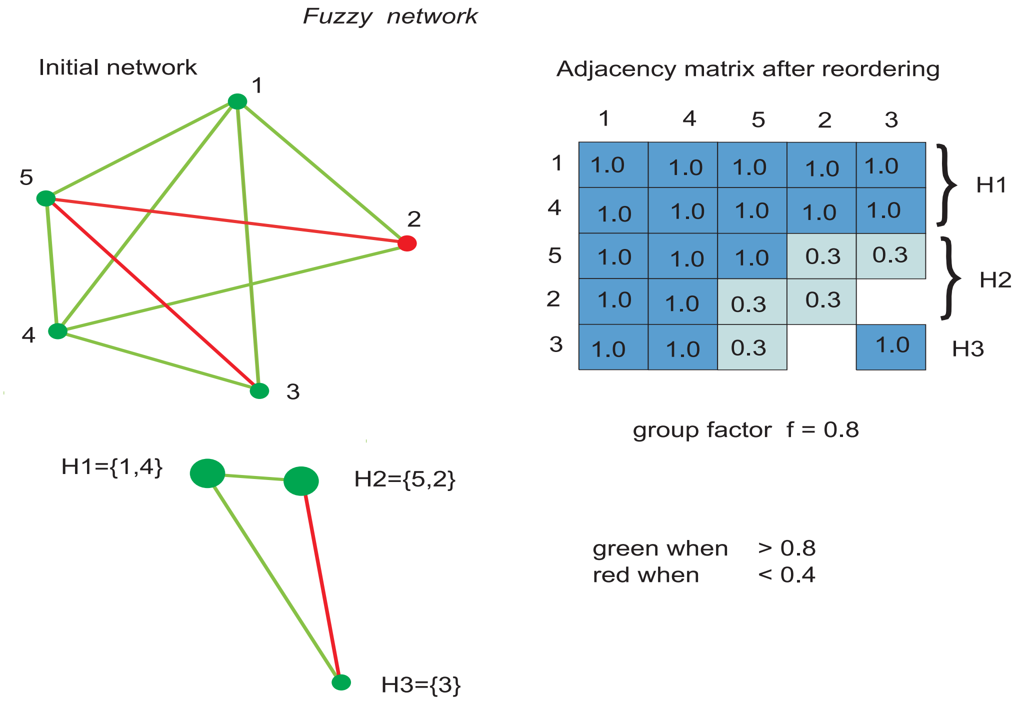
In this section we discuss briefly the issues around the propagation of uncertainty across a network, i.e. traversing nodes and edges. This is an important matter in a network in which nodes and links can appear and disappear with time. In such networks, the uncertainty of an observation increases with time since the observation. It is further complicated by the propagation of uncertainties and the dynamic interactions across the whole network. It is interesting to note that it is believed that many network types (e.g. social networks) are in general able to re-group themselves in response to varying degrees of structural changes or uncertainties to achieve stablity.

4.2.3.1 Propagation of Edge Uncertainties

In this section we will discuss the propagation of uncertainties or fuzziness of the edges in a network. Many possible approaches can address this problem, but among them a possible strategy is to follow the concept of how the usual union operator in fuzzy set theory is constructed. Here, the maximum value is selected, see for example page 50 in [13]. This means that when computing the strength of edge $r_{H1,H2}$ in Figure 4-11, for example, one should select the strongest link that connects a sub-node in hypernode $H1$ to a sub-node of hypernode $H2$. Node 4 is a member of $H1$, node 3 is a member of $H2$ and the link between nodes 3 and 4 has membership value $\mu(r_{3,4})=1$. Therefore, the edge $r_{H1,H2}$ has membership value 1.

4.2.3.2 Propagation of Node Uuncertainties

In a network, fuzziness applies as much to the nodes as to the edges; Figure 4-12 shows a network with a fuzzy node, in this example node 2 which is represented in red. The nodes are represented by the diagonal of the adjacency matrix. The motivation for introducing fuzziness of a node can be related to the knowledge about the uncertainty about the existence of a node. In the case considered, there are two red links in the network. The resulting reordered adjacency matrix and the corresponding network of hypernodes is shown at the lower part of the figure. Here, the edge $r_{H2,H3}$, for example, is red since there exist no strong links between the sub-nodes of $H2$ and $H3$. A property of the algorithm is that the uncertain node is mapped to hypernode $H2$. Since $H2$ is composed of a certain and an uncertain node, the maximum membership principle leads to the result that hypernode $H2$ is a certain node.



	1	4	5	2	3	
1	1.0	1.0	1.0	1.0	1.0	H1
4	1.0	1.0	1.0	1.0	1.0	
5	1.0	1.0	1.0	0.3	0.3	H2
2	1.0	1.0	0.3	0.3		
3	1.0	1.0	0.3		1.0	H3

Figure 4-12: The Weighted (Fuzzy) Network in Figure 4-6 is Modified by Introducing Edge $r_{2,5}$ as an Uncertain Edge and Node 2 as an Uncertain Node. The membership values are shown in the adjacency matrix. The group factor f is 0.8.

4.3 DEMONSTRATIONS OF THE ALGORITHM

4.3.1 Hypernodes Generated from a Book Example

Figure 4-13 illustrates the hypernode algorithm on a fuzzy relation used in Pedrycz and Gomide [15], page 88. The fuzzy relation is visualised by using the width of the lines to symbolize the strength of the edges. Here, the relation is shown at the 1st level and no aggregation of nodes is applied. In our case a network of hypernodes is constructed and the algorithms previously presented are used. The network is generated by applying the group factor $f = 0.6$. The figure shows that at the 2nd level there are four hypernodes connected with yellow or red links. At the 3rd level the sub-nodes are mapped to two hypernodes which are connected with a yellow link, i.e. a medium strong link. The traffic light symbology is again used to visualise the strength of the links. The hypernode approach creates a hierachical structure or layer of networks from the orginal unstructured network and provides a means to examine the structure of the network from its highest level of abstraction to its lowest level of detail information.

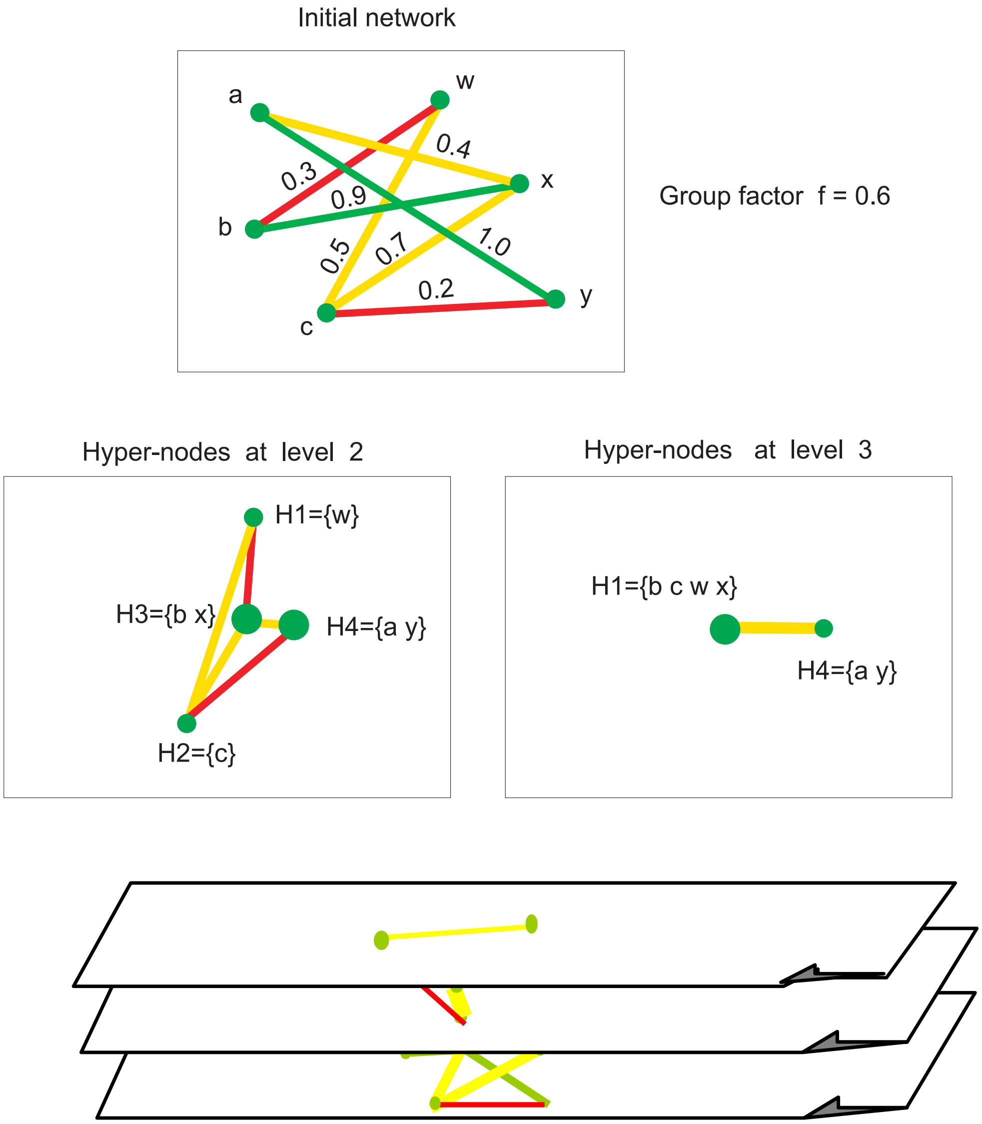


Figure 4-13: Hypernodes Generated from a Fuzzy Relation Example in [15], page 88.

4.3.2 A Geometrical Example

The point map in Figure 4-14 is used to generate a weighted network R, which is also shown as an adjacency matrix in the figure. The strength of the links of R is computed on the basis of the Euclidean distance between the points. Let $m(i,j)$ denote the distance between any two points i and j of the point set. The membership value $\mu(r_{i,j})$ is computed as a normalized value of $m(\cdot)$ as:

$$\mu(r_{i,j}) = \frac{m(i,j)}{\max[m(\cdot) \mid for\ all\ m(\cdot)\ in\ R]}. \tag{3}$$

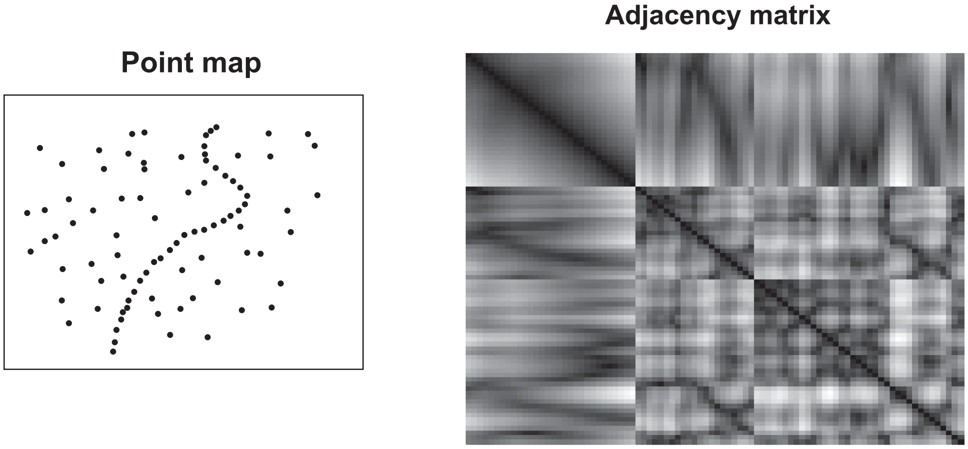


Figure 4-14: The Point Map Used to Generate a Weighted Network.

The adjacency matrix in Figure 4-14 is reordered as shown in Figure 4-15. By applying the algorithm in a recursive manner, hierarchies of adjacency matrices, i.e. hierarchies of hypernodes, can be constructed. The group factor is set to 0.9. The corresponding groups of points, i.e. point clusters, are shown in Figure 4-16. The metric selected in Equation 3 has the property that points located close to each other will be grouped together.

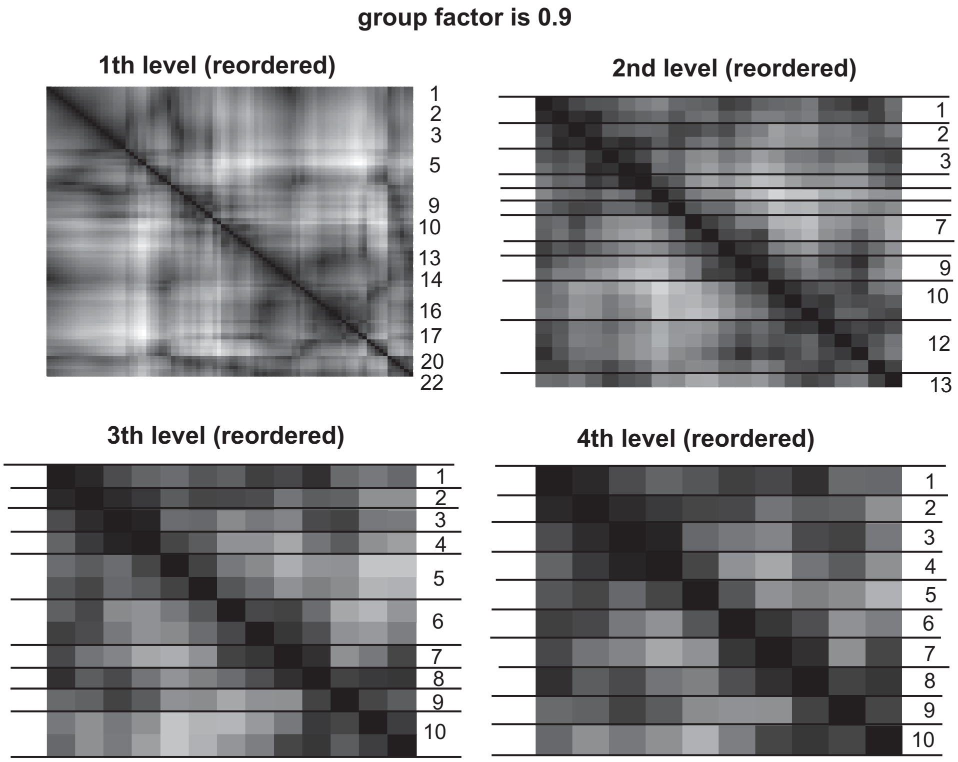


Figure 4-15: Hierarchies of Adjacency Matrices Constructed from Reordering the Adjacency Matrix in Figure 4-14. The applied group factor is 0.9.

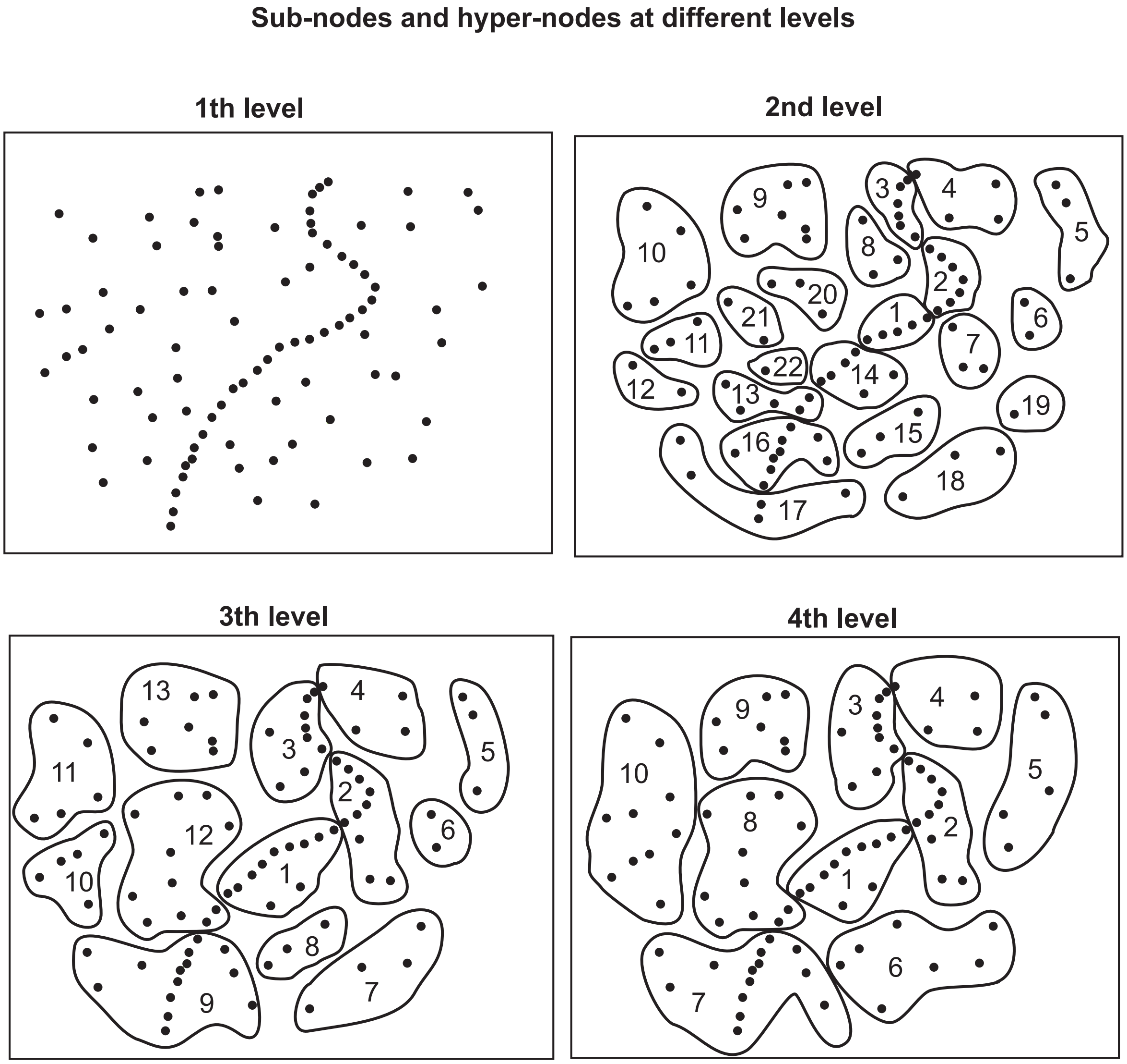


Figure 4-16: Point Clusters Corresponding to the Hierarchies of Adjacency Matrices in Figure 4-15. The identifiers of the clusters are the group numbers of the adjacency matrices.

The number of points in the initial point set, i.e. 1st level , is 88. At the 2nd level there are 22 point clusters. The next lower level represents a reduction of the number of clusters from 22 to 13, and the final level has 10 clusters.

The hypernodes at the 4th leve are shown in Figure 4-17 together with the original network. The large number of edges in the original network obliterates the visual clarity of the image. Therefore, from this cluttered image there is no means to understand the structure of the network, let alone being able to answer questions such as whether the network is separated into disjoint components. In the hypernode image there are only 10 hypernodes and the visual separation of the components of the network is clear and well presented. From this display it is clear that the network is made up of a single connected component.

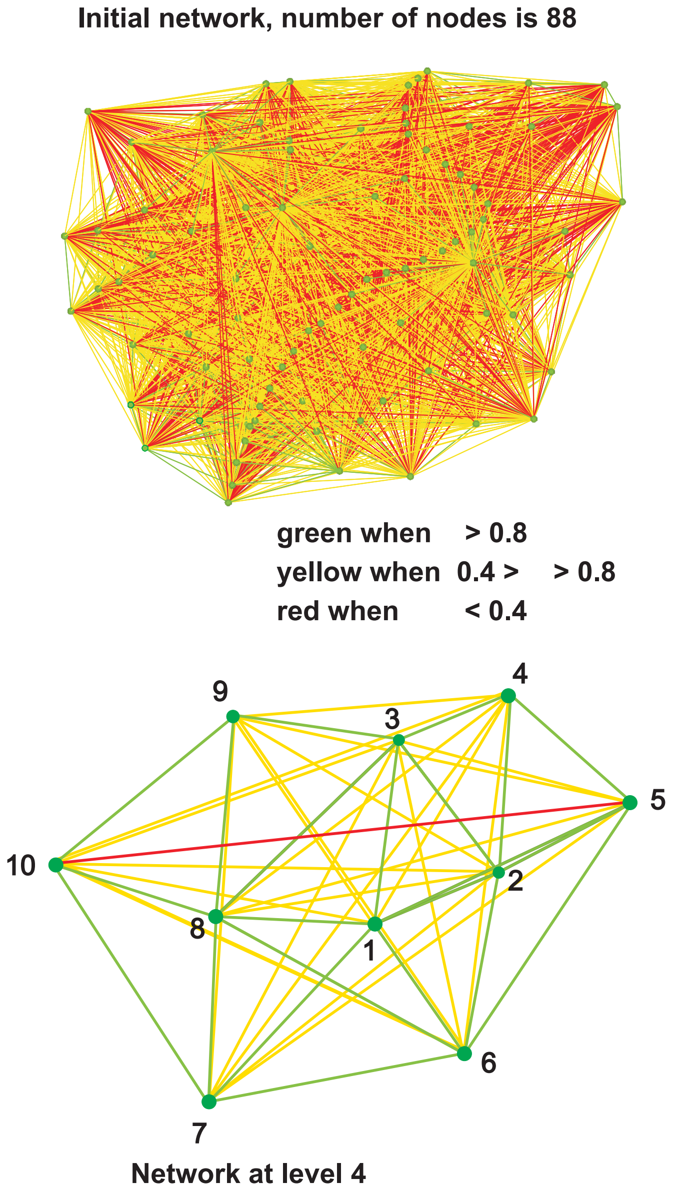


Figure 4-17: Network of Hypernodes Corresponding to the Adjacency Matrix at the 4th Level in Figure 4-16. The initial network is also shown. The colour coding shows the weight (fuzziness) of the edges of the network. Green colour means a strong or certain link, yellow means a medium strength link and red means a weak or uncertain link.

The group factor f controls the clustering of the nodes. In order to demonstrate the effect of the group factor, hierarchies of hypernodes are computed for different values of f. The result of this computation is shown in Figure 4-18. The figure shows how the number of levels and the number of hypernodes at the top level of the hierarchy depend on f. When f is greater than 0.9, the number of point clusters at the top level increases rapidly with increasing value of f. When f is less than 0.81, the hypernode at the top level contains all the nodes of the original network. The number of levels of the hierarchy has maximum when $f = 0.87$.

The group factor is a case and application dependent parameter which plays a significant role in the determination of the resulting hypernode structure.

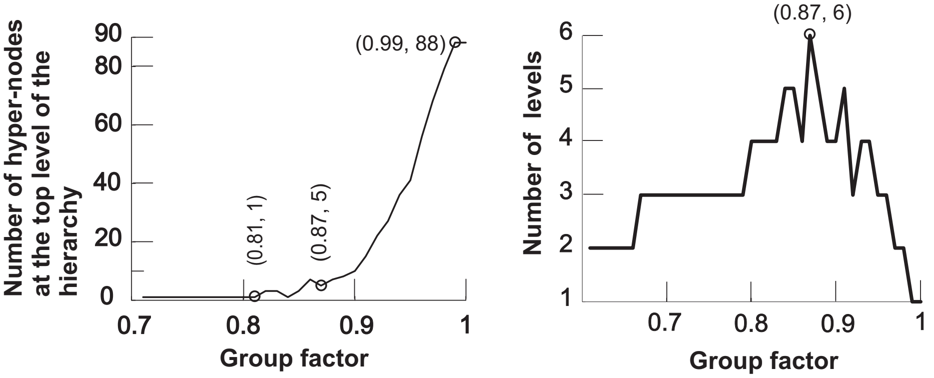


Figure 4-18: Hypernodes Computed for Different Group-Factors of the Network in Figure 4-17. The two graphs show how the number of hypernodes at the top level of the hierarchy and the number of levels depend on the group factor.

4.4 DISCUSSION

In this chapter we describe a generalised approach to constructing networks of hypernodes. However, in this approach the only criterion for the construction of the networks of hypernodes is the number and strenght of the links. Other criteria can be used, as discussed in Annex E.

The group factor f is crucial to the construction of the hypernode structure; it is therefore important for the user to be able to interactively manipulate, study, visualise and understand the effect of different settings of f. One approach is to find a group factor that generates a number of hypernodes less than a certain threshold, but there are many possible approaches.

The hypernode can be extended to provide detailed information about its sub-nodes and their connections so that the user can zoom into the hypernodes to access information about the underlying nodes and their connections.

The concept of weighted networks can be utilized to model the uncertainty of the different components of the network, but the simple way the proposed algorithm propagates the membership value of sub-components to hyper-components presupposes that the maximum criterion is a suitable model for the propagation of the uncertainty.

The visual representation of a weighted network should map the fuzziness of the network to an appropriate visual variable. Bertin [3] argues that quantitative information should be mapped to visual variables that are able to reflect the ordering of the data. According to Bertin the appropriate visual variables for this purpose are size and gray value. Since hue is a qualitative property of colours, Bertin argues that hue cannot offer an intuitive connection to a quantitative information variable. However, colour hue is a very selective visual variable, for example the traffic light is an intuitive representation and offers good natural separation between the three groups, see for example Figure 4-20. A study into human visual perception of uncertainties could provide a more in depth understanding, but such a study is beyond the scope of the current work.

An alternative visual coding is shown in Figure 4-19. Here, the visual variable gray value is applied. A comparison between the traffic light representation and the gray value principle must consider different parameters like line width, background colour, the actual definition of the colours, etc. For example, should the yellow colour be moved a little bit in the direction of orange in order to optimize its visibility? Since colour hue is a very selective visual variable, we can probably assume that the traffic light representation will give a clearer separation of three classes than the gray scale. What degree of accuracy is required to represent the uncertainties for instance in the yellow colour and what benefits or hinderance to the user if for example the yellow colour is further divided into light, medium, darkish or orangy yellow. Detailed analysis of this topic is outside the scope of the present chapter.

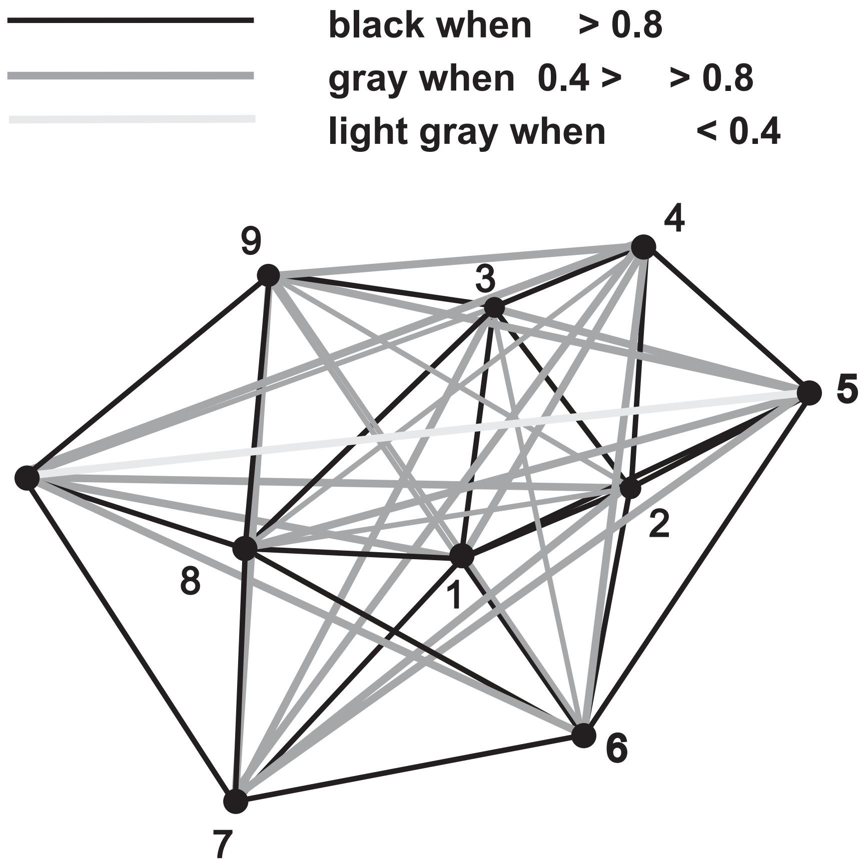


Figure 4-19: Alternative Colour Coding of the Network in Figure 4-17.

Figure 4-20 demonstrates how a weighted network can be divided into several windows, one for each class of membership values. From window A, for example, it can easily be seen that all the nodes are connected with strong links; this can also be said for the medium strong links in window B. From window C it is clear that only two nodes are connected with a weak edge. The separation of the network into three windows has the advantage that the different classes of strength (uncertainty) can be studied separately. In window D all the edges are shown. Here, it can be a little bit harder than in window A to answer the question which nodes are connected with strong links, for example. This example shows the benefits of the traffic light colour sheme in easing the understanding and extraction of information from the network. Alternatively, we can present this as a layered network, Figure 4-12.

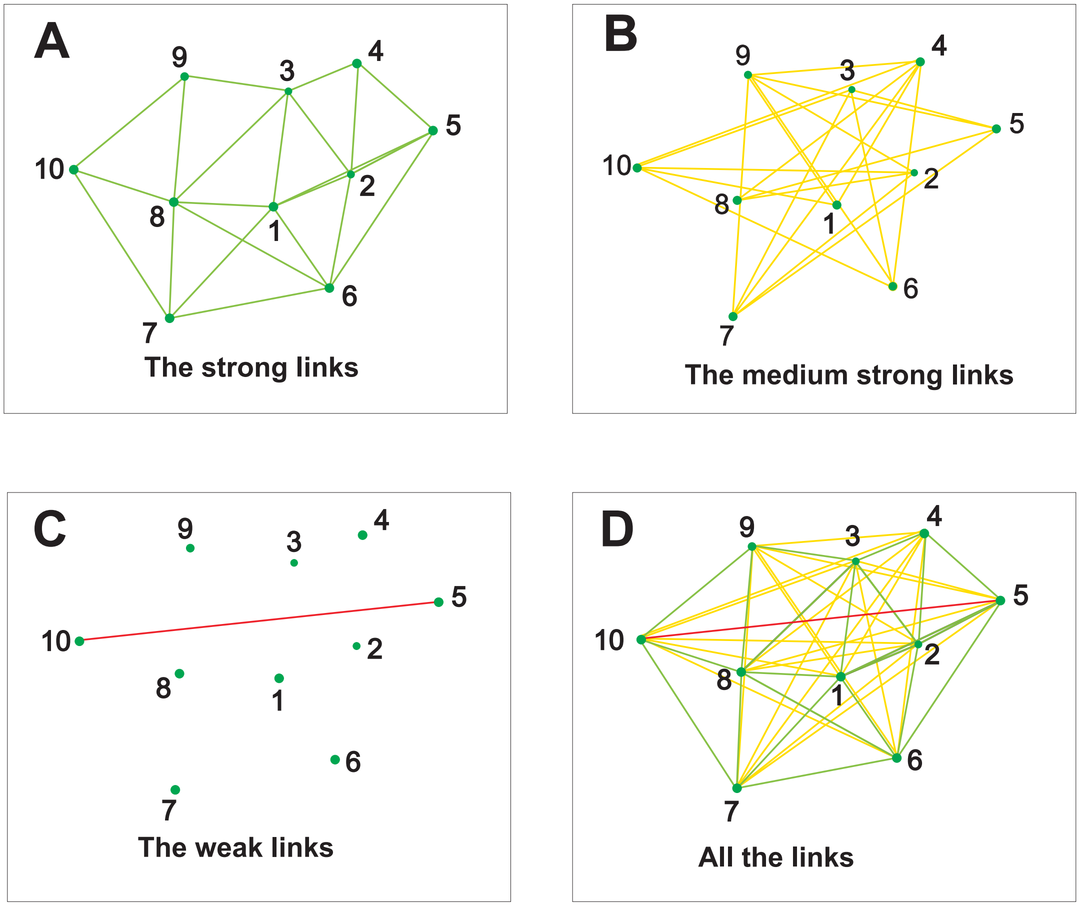


Figure 4-20: The Network of Hypernodes in Figure 4-17 Divided into Three Windows, i.e. One Window for Each of the Three Classes of Links.

When the network is embedded in a geographical space, the mapping of the network to a position in the plane at the first level is trivial, i.e. the geographical co-ordinates can be used to locate the nodes. At the generalized levels of a network the situation is different; here, the hypernodes are abstract nodes and have no natural geographical location. The question is therefore where to position the hypernodes. In the examples presented, the position of a hypernode is computed as the average position of its sub-nodes.

Fabrikant et al. [9] use the term spatialization when mapping non-spatial data to an information display. Information spatialization is inspired by the analogy that the strength of relatedness in the data space should be mapped to neighbourhood in the information display, such that semantically similar nodes are placed closer to one other than less similar ones. An empirical study suggests that the distance-similarity metaphor applies to network spatializations by equating metric distance along network lines to similarity. They also find that line size, colour value and hue, modify the distance-similarity metaphor in subtle yet logical ways. An implementation of the spatialization principle [9] is not straight forward, since the mapping of a weighted relation to a two-dimensional plane cannot always guarantee that the strength of the relations is mapped to neighbourhood in the plane.

Distinction can be made between visual communication and visual exploration [14]. Visual communication deals with how to represent results of different kinds of analysis, i.e. whenthe message is well defined.

The other view is covered by visual exploration; in this case the message is not well defined and the analysis of structure or detection of important information is left to the information receiver. Thgis implies that exploration is permitted by the provision of various tools. A full discussion of this topic is not in the scope of this chapter, but see Chapter 2 and Annex B for a more extended discussion of the differences among monitoring, searchng and exploring tasks. From the exploration point of view we recommend that the analyst should be provided with tools to examine/assess the effect of:

- The group factor;
- The thresholding functions to visualise edges in a certain interval of membership values; and
- The selection of visual variables to illustrate the degree of strength of the different edges, e.g. traffic light symbology, gray scale or line size.

The time complexity of the proposed implementation grows as order N^2. Therefore, when the size of the network exceeds a certain threshold, for example 1000 nodes, strategies to reduce the computing time must be considered.

4.5 CONCLUSIONS

The hypernode algorithm has been described and demonstrated. The experimental results shown are based on the proposed algorithm. The algorithm discussed considers the strength of the relations, it can be used to construct networks of hypernode of weighted (such as fuzzy) as well as crisp relations. The crisp case comes out as the special case where the membership values are either 0 or 1. The introduction of membership values allows uncertainty of crisp networks to be studied.

Hypernodes can be constructed at different levels, i.e. the degree of generalization or abstraction increases with the number of iterations the algorithm. The process terminates either when there is one single hypernode is achieved at a particular level or when the selected group factor does not allow further grouping of the nodes. The mapping of the nodes to hypernodes is a many to one mapping, i.e. one sub-node can be mapped to only one hypernode at a certain level.

The algorithm allows networks to be studied at different levels of abstraction. In that way a high level understanding of the network can be obtained. Indeed the hypernode transforms a flat network into a hierarchical structure that reveals the underlying structure and pattern of the network in an effective and intuitive manner.

The hypernodes and the edges between them can be utilized to study the effect on the network when groups of nodes or groups of edges are eliminated from the network. For example, one can ask what happens to the network when a certain hypernode is destroyed. In this way information about the vulnerability of the network can be identified so that strategy can be developed to strengthen own force or weaken enemy.

The geometrical example indicates that the algorithm can be applied to the clustering of points in a geographical space. This can be utilized in, for example, cartographic generalization.

We have shown how uncertainty about the edges and nodes propagates in a network. However, it is important to understand how the uncertainties of these individual edges and nodes combine and affect the uncertainty in the entire network. Also, how do these uncertainties propagate through for the network to restore its stable state?

A topic for further research is to apply the algorithm to real situations and study how it can be adapted to meet the needs of different use cases. The propagation of the weights or nodes and edges to higher levels or the hierarchy, will be adressed and further developed.

4.6 REFERENCES

[1] Berge, C., “Graphs and Hypergraphs”, North-Holland, Amsterdam-London, 1973.

[2] Berge, C., “Hypergraphs: Combinatorics in Finite Sets”, Elsevier, New York, 1989.

[3] Bertin, J., “Graphics and Graphic Information Processing”, Walter de Gruyter, Berlin, New York, 1981, 273 pages.

[4] Bjørke, J.T., “Framework for Entropy-Based Map Evaluation”, *Cartography and Geographic Information Systems*, 23(2):78-95, 1996.

[5] Bjørke, J.T. and Smith, B., “Seriation: An Implementation and Case Study”, *Computers, Environments and Urban Systems*, 20(6):427-438, 1996.

[6] Bjørke, J.T., “Reduction of Complexity: An Aspect of Network Visualization”, *RTO-MP-IST-063 Visualising Network Information, NATO Conference, Copenhagen, Denmark*, pp. 9-1 to 9-10, 2006. NATO, ISBNs 92-837-1156-4 or 978-92-837-1156-8.

[7] Bjørke, J.T. and Varga, M.J., “Algorithm to Construct Fuzzy Hyper-Networks”, Presentation at the *NATO IST-059/RTG-025 Network of Experts (N/X) Workshop: Network Analysis and Visualisation for Simulation and Prediction*, El Segundo, California, 2007.

[8] Börner, K., Sanyal, S. and Vespignani, A., “Network Science”, In Blaise Cronin, Editors, *Annual Review of Information Science and Technology*, Chapter 12, pp. 537-607, Information Today, Inc./American Society for Information Science and Technology, Medford, NJ, 2007.

[9] Fabrikant, S.I., Montello, D.R., Ruocco, M.and Middleton, R.S., “The Distance-Similarity Metaphor in Network-Display Spatializations”, *Cartography and Geographic Information Science*, 31(4):237-252, 2004.

[10] Flake, G.W., Tarjan, R.E. and Tsioutsiouliklis, K., “Graph Clustering and Minimum Cut Trees”, *Internet Mathematics*, 1(4):385-408, 2004.

[11] Huang, X. and Lai, W., “Clustering Graphs for Visualization Via Node Similarities”, *Journal of Visual Language & Computing*, 17:225-253, 2006.

[12] NATO IST-059/RTG-025, “A Framework for Network Visualisation”, *RTO Technical Report*, NATO, http://www.rta.nato.int/, 2007.

[13] Klir, G.J. and Yuan, B., “Fuzzy Sets and Fuzzy Logic: Theory and Applications”, Prentice Hall P T R, New Yersey, 1995, 574 pages.

[14] MacEachren, A.M., “Visualization in Modern Cartography: Setting the Agenda”, In MacEachren, Alan M. and Taylor, D.R. Fraser, Editors, *Visualization in Modern Cartography*, pp. 1-12. Pergamon, Oxford, 1994.

[15] Pedrycz, W. and Gomide, F., "An Introduction to Fuzzy Sets: Analysis and Design", The MIT Press, Cambridge, Massachusetts, 1998, 465 pages.

[16] Siirtola, H. and Mäkinen, E., "Constructing and Reconstructing the Reorderable Matrix", *Information Visualization*, 4(1):32-48, 2005.

[17] Varga, M.J. and Bjørke, J.T., "Hyper-Networks with Uncertainties", Presentation at the *NATO IST-059/ RTG-025 Network of Experts (N/X) Workshop: Network Analysis and Visualisation for Simulation and Prediction*, El Segundo, California, 2007.

[18] Borkowski A., Grabska E. and Szuba, J., "Visualisation of Graphs in ArchiCAD", *Lecture Notes in Computer Science*, Vol. 1774, 1999, pp. 241-246.

[19] Lee-Kwang, H. and Cho, C.H., "Hierarchical Reduction and Partition of Hypergraph", *IEEE Transaction on Systems, Man, and Cybernetics – Part B:Cybernetics*, Vol. 26, Issue 2, April 1996.

[20] Bjørke, J.T. and Varga, M.J., "Weight (Uncertainty) Propagation in Networks of Hypernodes", *NATO Network-of-Experts Workshop on Visualisation of Network Dynamics*, UK, November 2008.

Chapter 5 – FRAMEWORK/SURVEY INTEGRATION

The IST-059 Framework for network visualisation is discussed in Chapter 2 and Annex B, and the Survey of network visualisation applications and tools is treated in Chapter 3 and Annex F. Each is potentially useful in its own right, but IST-059 decided during the 2006 IST-063 workshop in Copenhagen [1] that integrating them might well provide added value. The framework allows the user to specify the problem in a semi-formal manner that should suggest what kind of representations and display techniques might be most appropriate. The survey lists the attributes of a large selection of tools for displaying networks and their properties. It therefore seems natural that the framework should be developed to become useable as a front-end for querying the survey database.

5.1 FRAMEWORK AND SURVEY TOGETHER

During the life of IST-059, a start was made on a conceptual design and the description of a procedure that can be employed manually by the end-user. No work was done to implement the framework in a computationally useful form, and therefore no work was done toward integrating the framework with the survey beyond the initial conceptual design. This chapter outlines the initial design and the concept for further development.

5.2 END-USER’S VIEWPOINT

When an end-user studies a network, it is usually not because the network is interesting in itself, but because the user believes that some real-world task might benefit from information embodied in the network. Among the many software tools that have been developed for extracting and displaying properties of networks, it is quite likely that one or more would be useful for the task at hand. However, it is also likely that no tool was developed for this user’s particular current task, and it is also quite likely that a non-specialist user will be unaware of the full range of available software tools. Most tools are developed either to solve a particular real-world problem, or are developed for studies into more general network properties. In either case, unless the user is following a well-trodden path, it is unlikely that any particular tool was designed with the user’s current task in mind. A suitable tool, if one exists, must be sought among the many that have been developed for other purposes.

The IST-059 Survey was developed independently of the Framework, with a view to providing a partially populated database structure that an end-user might query to discover tools that might offer particular analytic or display algorithms. However, to use the Survey effectively, the end-user must have some specialized knowledge of network analysis and display technology, at least enough to know what algorithms or display techniques or data formats to use as search query terms.

The IST-059 Framework did not concern itself with specific applications or tools. It was intended as a framework within which the properties of networks and tasks that involve networks can be described in a uniform manner. The concept was that the Framework might be used to evaluate the strengths and weaknesses of particular software tools in relation to the requirements of the user’s task, and to suggest what characteristics of display design would probably be useful for the particular task.

The task side of the Framework was based around the earlier work of IST-013 and IST-021, who developed the functional “VisTG Reference Model” for visualisation (Annex H), together with the RM-Vis descriptive framework developed by TTCP C3I AGVis. If an end-user knew of a particular software tool and its properties, as well as having a good understanding of the real-world task, the Framework should help the user to determine the usefulness of the tool for the task.

Midway through the life of IST-059, it became apparent that the Framework might well be integrated with the Survey to make a tool more powerful than either. At the IST-063 Workshop in Copenhagen, one of the working groups studied how this might be done, and reported a development concept for the integration. The Framework-Survey Integration Group proposed that the process be structured around a worksheet, which could be paper-based or might be embodied in software. The worksheet suggests a workflow that follows naturally from Figure 5-1 (reproduced from part of Figure 2-5).

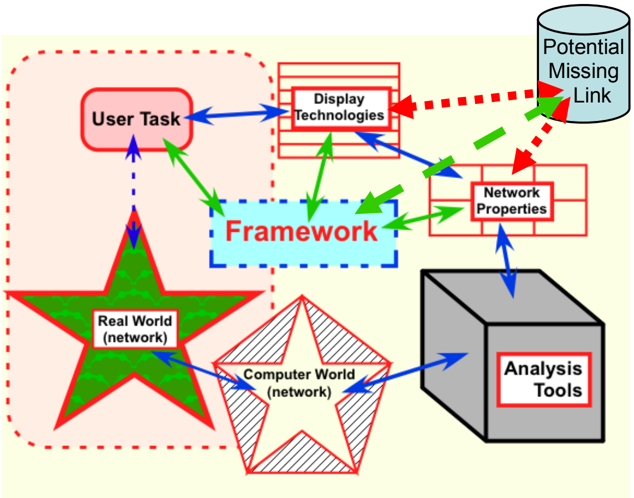


Figure 5-1: The Place of the Framework and Survey in the User's Task Flow.

In the absence of the Framework and the Survey, the user's conceptual workflow is shown by the blue arrows in Figure 5-1, as follows:

1) The user starts by analyzing what task-relevant information could potentially be gained from studying some network in the real world.
2) Some of the data that define the real-world network are abstracted into the dataset stored in the computer.
3) The dataset in the computer probably has some properties relevant to the user's task, and these properties may be extracted by the use of selected analysis tools.
4) For the user to access the extracted properties, a display and interaction technology must be chosen.
5) Finally, a display presentation must be produced that assists the user to visualise the application of those properties to the real world task.

Many applications for working with network data exist. Most incorporate all five stages of this workflow. Some are monolithic; when presented with the data, they produce a screen display representing the results of their processing. Others may allow the user choices at some or all of the stages. It is up to the user to select the application, and to choose which choices to make if the selected application permits choice at any stage. The user's ability to select the application, and to make effective choices within the application, depends on the background knowledge of the individual.

It is most unlikely that any but a specialist will know enough about more than a handful of the hundreds of possibilities to be able to make an informed choice. The Framework and Survey are intended to ease this task. The Survey is intended to describe available applications in such a way that the user can discover which ones might be useful for the task, whereas the Framework is intended to aid the user in developing queries to the Survey, as well as to guide the stages in the regular workflow.

5.3 FRAMEWORK WORKFLOW

The workflow is slightly different when the Framework is used as a process. Conceptually, the same things need to be established: what about the real-world task involves a network, what data about the network are available in the computer's database, what properties of the network would be relevant to the real-world task, what algorithms can extract those properties, and how the results of executing those algorithms should be displayed.

The Framework workflow is currently based around a worksheet on which the user answers several series of questions. Examples of the use of this worksheet are shown in Chapter 6. The worksheet is a first draft, and is expected to be amended after it is used to address real problems. As matters stand at the end of the life of IST-059, it has been used for four somewhat artificial use cases in different domains. An effective implementation of the Framework would replace the worksheet by a software structure that uses the answers to the questions to create a query interface to the Survey. The development of such a structure is among the recommendations to be passed to IST-085.

The first-draft Framework has two groups of question sets. Questions in the first group mostly require textual answers, whereas questions in the second group mostly involve checking off tick-boxes. There are four question sets in the first group, the first of which contains questions about the real-world task and the modes of perception (Control/Monitor, Search, Explore, Alert) involved. The answers to these questions serve two purposes: firstly, thinking about how to answer may help the user to clarify just what the problem is, and secondly, the answers affect the appropriate choices of display and interaction technology for the final stage of the normal workflow pattern. The knowledge required to answer these questions is of the real-world task, not of networks and their properties.

The second set of questions is about the network of interest. The user is asked about the categories of nodes and links, and about traffic over the links and any timing effects related to traffic. To answer these questions requires more knowledge of networks than do the questions of the first set, but still the answers are based more on the real-world requirements of the user than on the abstract properties of networks.

The third set of questions define the characteristics of any embedding fields (relevant context), which require no specialized knowledge of networks. These are followed by questions about the measures that will be used in addressing the network aspects of the real-world problem, questions which do require expertise in network analysis.

Finally, in the fourth question set of this first group, the user is asked about the data source – questions such as the types of data, their reliability, and whether all the data are available at the start of the analysis, whether more can be sought to fill gaps, or whether the data are streamed in real time.

Whereas the first group of question sets in the first draft worksheet requests textual answers, the second group is answered by ticking off characteristics that apply. They complement the first group, and should be consistent with it. The questions in the second group are more technical in nature, and some may be problematic for a user whose expertise is with the real-world problem domain. They concern the task dynamics and interactivity, the perceptual modes, the technical nature of the network, and the data characteristics following the taxonomy of Annex B, Section B.3.1.2.

5.4 FRAMEWORK-SURVEY INTEGRATION CONCEPT

Let us imagine that the worksheet has been refined through a series of tests using real problems, and that it has been replaced by a Web-based user interface backed by software capable of interpreting the answers in a form corresponding to the network and display characteristics described in Annex B. The workflow described in Figure 5-1 shows the Framework interacting with the Survey database and with three stages of the normal workflow: Network Properties, Display Technologies, and User task.

These interactions can be viewed a little differently if the Framework is taken to serve as an interface between the processes in the normal workflow and the Survey database, filtered according to the answers provided by the user to the sets of questions. From this viewpoint, as shown in Figure 5-2, there are four stages, loosely sequential: Network Properties, Data Type, Display Requirements, and Display Design. Each of these stages provides a different filter that can be used to generate a query to the Survey database. Those tools that survive the filtering process form a pool of software that the user might consider further.

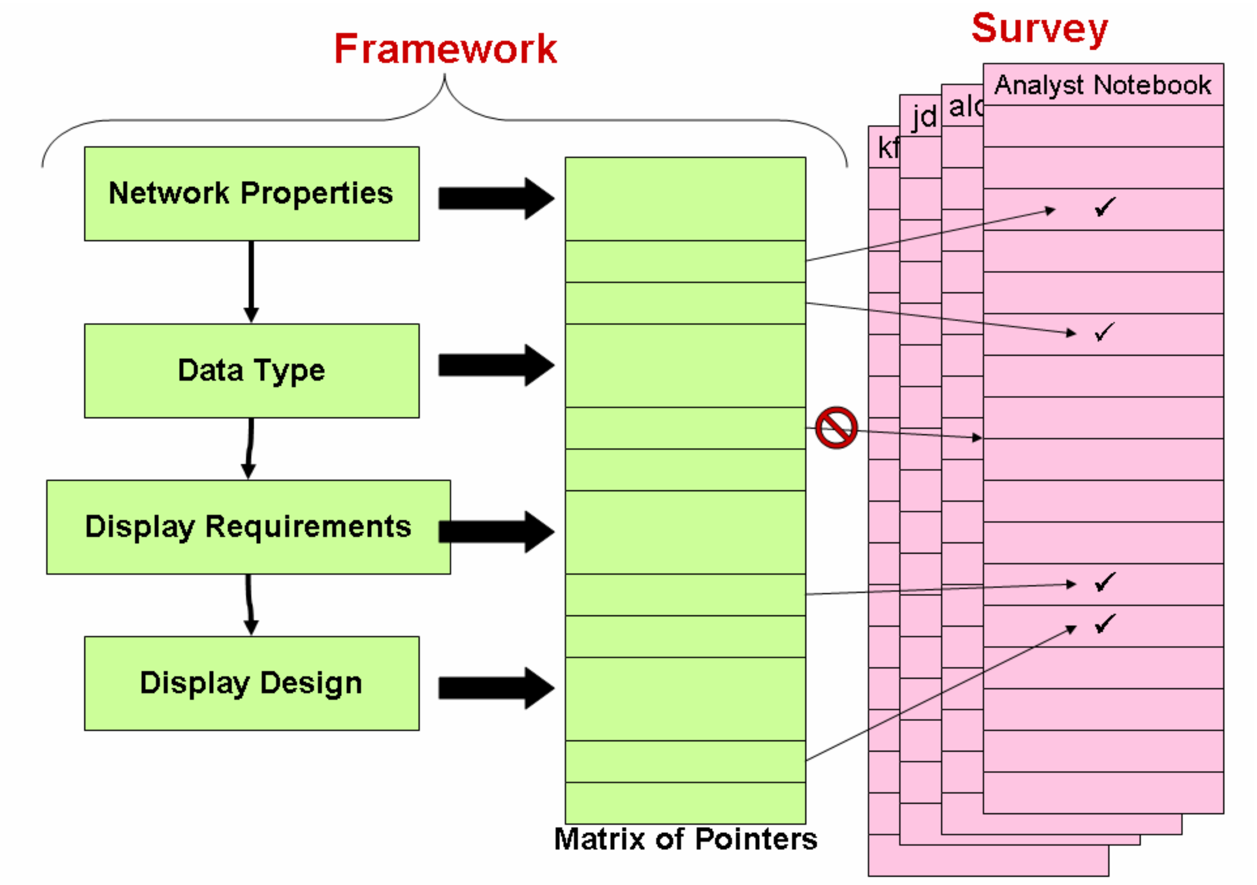


Figure 5-2: The Conceptual Flow of the Framework and Survey.

The Framework cannot query the Survey until the User has specified the real world task, at least to the extent of defining what kinds of network properties might be useful to visualise: properties of the network as a whole, or of local parts, static properties or dynamic developments, monomodal or multimodal aspects, and so forth. In other words, before the Framework can be effective, the user must have some idea of what would represent a successful view of the network. To this end, the workflow begins with the user filling out an actual worksheet, or mentally doing the equivalent.

The preliminary questions relate to the task and to the data. Some of them may be difficult for the user to answer, but they form the basis of the kinds of query that will need to be submitted to the Survey database if it is to report applications suitable to the user's task. The Framework deals not with the task and the source data, but with the relationship of these to the applications and displays that might help the user.

When the user has thought through the real-world problem, certain things should have been clarified. For example, the user should have a good idea about the kind of network or networks involved, and about some of the network properties that could help his or her understanding. If that is the case, there is no point in considering an application that cannot deal with the known kind of network, or that cannot extract and display the interesting properties.

5.5 USING THE FRAMEWORK

Before the user can effectively deploy the results of the queries that the Framework has used to filter the available software tools, several other considerations strongly influence the usability of any tool, as suggested in Figure 5-3, developed by the Integration Working Group at the 2007 N/X Workshop in El Segundo, CA, USA. Figure 5-3 is a proposal for the central portion of the VisTG Reference Model View of the Framework, labelled "Human Factors Engineering" in Figure B-12.

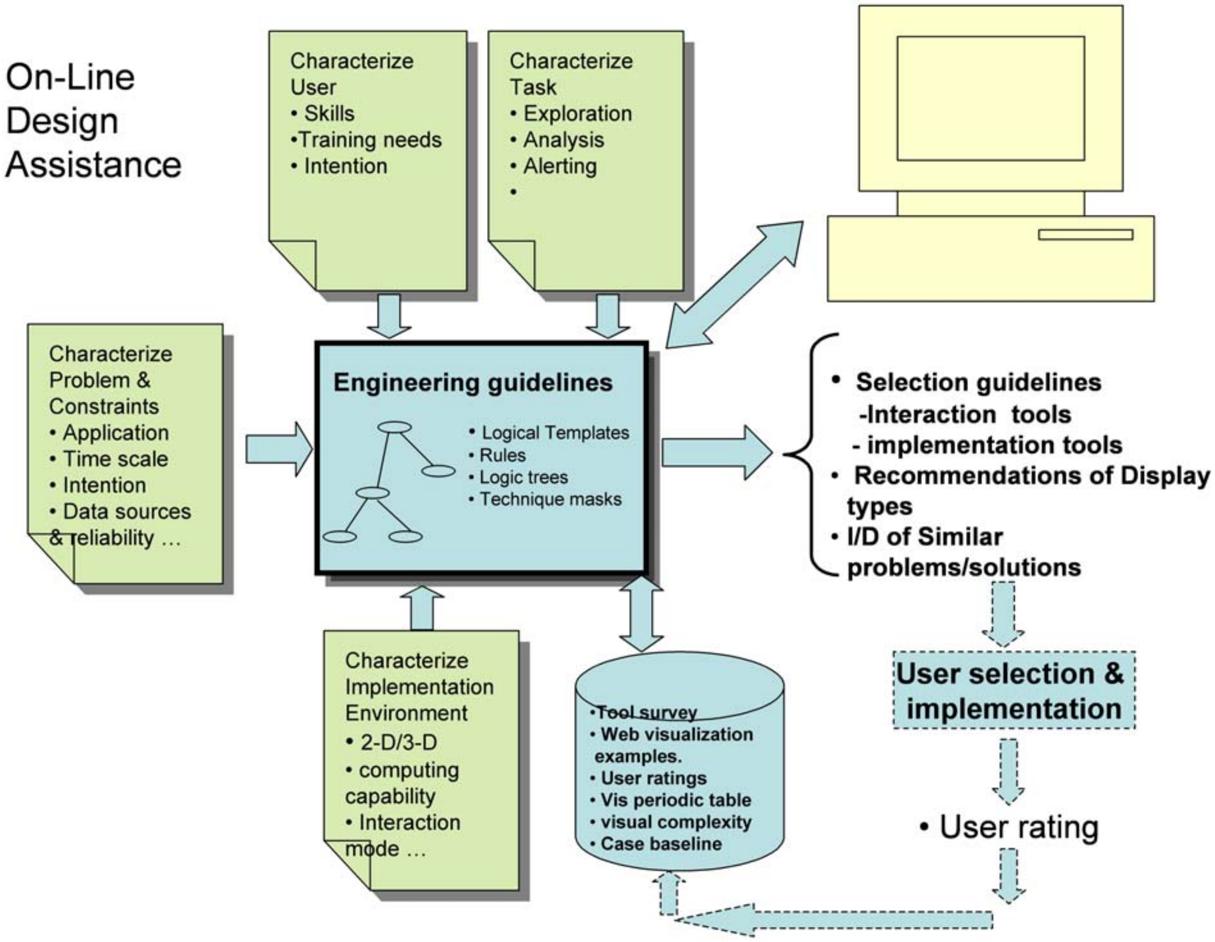


Figure 5-3: Further Considerations in Using a Future Implementation of the Framework, Centred around the Final Display.

Figure 5-3 offers a contrasting viewpoint to that of the Framework worksheet. Whereas the Framework worksheet is focused on what the user wants to achieve by understanding something about the network, Figure 5-3 is centred on Engineering Guidelines for display design. Within the conceptual workflow suggested in Figure 5-2, Figure 5-3 provides detail for the final stages: setting the display requirements and deciding on a display design.

Among the inputs to the central “Engineering Guidelines” box in Figure 5-3 are “Characterize User” and “Characterize Implementation Environment”. Neither was included in the draft worksheet, in part because no obvious way of characterizing a user presented itself. It is, however, one of the dimensions of the RM-Vis framework (Annex G) which is one of the sources for the IST-059 Framework. A revised worksheet draft should incorporate questions that would characterize the user’s relevant skills and limitations. The implementation environment often is a question of scale, but it can also make a considerable difference to the decisions as to the most appropriate type of display. An immersive interactive 3-D environment is very different from a hand-held portable device that might be carried by a soldier in the field. Either might be appropriate for some particular network task, but the modes of display suited to one would be far from optimal for the other.

5.6 FRAMEWORK VIEW OF THE SURVEY

If the Framework is to provide an interface through which the user can interrogate the Survey, two things are required:

1) The Survey database must incorporate the required information, meaning both that it contains the necessary data fields and that the data for particular applications has been properly entered, and
2) The Survey query interface must be structured appropriately.

The current (June 1, 2007) set of data fields for the Survey database is listed in Annex F. Since the database structure was developed independently of the Framework, mismatches exist between the Framework taxonomies and the Survey data fields, which are based directly on the RM-Vis framework.

A serious problem with the existing Survey structure is highlighted by the fact that for many applications the required information has not been entered. The application is shown as existing, and skeletal information about it has been entered, but no detail. This would not matter, were it not for the fact that structurally there is no distinction between "capability absent" and "capability not entered". Capabilities are represented by binary checkboxes that assert the presence of the capability, if its presence has been specifically noted. There is therefore no way for a Framework-based query interface to determine whether a particular application should be considered or excluded when a specific capability is required but the database does not show that the application has that capability.

To continue the discussion, assume that the database has been restructured to provide a ternary representation for the existence of a capability: "Present", "Absent", "Unspecified". The user has determined at least some of the attributes of the real-world network, and applications that consider all of the attributes explicitly are likely to be more useful than are ones that deal with them only implicitly or fail to deal with them. For example, if links in the real-world network are polyvalent and if some of the link components are characterized by having different traffic capacities, usage, and availabilities, then an application that treats all links as simply existing or not existing is unlikely to be much use.

The existing survey interface for entering application information has several different sections, which should be aligned with the Framework workflow. The complete set of attributes that can be entered for a specific application through the Web interface is listed in Annex F. These attributes are collected in several sections: Basic Information, Network, Analysis, Representation, Deployment, Acquisition, and References.

5.6.1 Survey Data and Further Possibilities

The elements in this section describe in broad terms what the application is intended for. It includes a cursory set of attributes from the RM-Vis four-axis descriptive model for visualisation applications. A more complete set for military applications, though still far from exhaustive, was provided by Vernik and Bouchard [2], and is shown by example in the following table.

Table 5-1: Domain Context Model

CATEGORY		ATTRIBUTE
Where	**Level of Command**	• Operational, Strategic, Tactical
	Environment	• Air, Land, Maritime, Joint, Littoral, Space, Urban
	Area	• Acquisition, Communications, Development, Engineering, Intelligence, Operations, Personnel, Plans, Requirements, Research, Training
	Scenario	• Humanitarian Assistance, Low/Medium/High Intensity Conflict, Peace Support, Special Ops
Who	**Role**	• COS, Commander, J2, J6, J7, J8, Intel Analyst, Logistics Officer, Ops Officer, Support Engineer
Why	**Activity**	• Analysis, Assess, Assign, Execute, Monitor, Plan, Report, Schedule, Track

It may not be necessary for the Survey database to record domain information in such detail. In fact, it would be likely to be counterproductive, if searches then failed to find a suitable application simply because it had been listed as being developed for a different domain context. It is important, therefore, to abstract from the different domain contexts those aspects that would help in assessing the degree to which usefulness in one context would suggest that an application would also be useful in another.

Vernik and Bouchard provide examples of viewpoints that might be taken onto network data in different domain contexts. These have implications for the kinds of display that might be expected to be useful.

Table 5-2: Example of Viewpoints

VIEWPOINT	DOMAIN CONTEXT	DESCRIPTIVE ASPECT
Monitor Belligerent Activities	**Activity**: Analyse, Assess, Monitor, Track **Area**: Communications, Intelligence **Environment**: Joint, Land, Urban **Level of Command**: Strategic, Tactical **Role**: HQ J2, Intel Analyst **Scenario**: Special Ops, Low Intensity Conflict	**Communications**: Email, Phone **Events**: Sequence **Finance**: Currency, Money **Geography**: Area, City, Country, Origin, Destination, Location, Maps **Identity**: Name, Sex **Information**: Document, File, Opinion **Movement**: Flight, Travel **Occupation**: Activity, Engagement **Organisation**: Unit People: Belligerent, Group, Organisation, Warlord **Relationships**: Degree, Enemy, Friend, Non-friend **State**: Alive, Dead **Time**: Age, Critical, Current, Date, Duration, Interval **Transportation**: Vehicle, Car
Assess Communication Network Robustness	**Activity**: Analyse, Assess **Area**: Communications **Environment**: ALL **Level of Command**: Tactical **Role**: Support Engineer **Scenario**: ALL	**Computer**: Hardware, Network **Geography**: Location, Maps, Latitude, Longitude **Telecommunication**: IP, Network, Parabolic-dish, Satellite **Usage**: Frequency
Team Building	**Activity**: Assign, Plan **Area**: Personnel, Training **Environment**: ALL **Level of Command**: ALL **Role**: Ops Officer **Scenario**: ALL	**Ability**: Skill **Assignment**: Mission, Order **Capacity**: Force **Events**: Scenario, Sequence **Identity**: Name, Sex **Occupation**: Activity, Engagement, Function, Jobs, Responsibility, Task, Work **Organisation**: Unit **People**: Group, Organisation, Person, Player, Soldier **Relationships**: Friend **State**: Ready, Standby, Not-Ready, Morale **Time**: Age, Deadline, Duration, Priority

Table 5-3: Viewpoint for Analyst's Notebook Geography View

DOMAIN CONTEXT	DESCRIPTIVE ASPECT
Activity: Analyse, Assess, Report, Schedule **Area**: Intelligence **Environment**: Joint, Land **Level of Command**: Strategic, Tactical **Role**: HQ J2, Intel Analyst **Scenario**: Special Ops, Low Intensity Conflict	**Communications**: Email, Phone **Family**: Brother, Sister **Finance**: Account, Money, Transfer **Geography**: Area, City, Country, Location, Maps **Identity**: Name, Sex, Flag, Nationality **Movement**: Flight, Travel **People**: Actor, Group **Possession**: Holder **Relationships**: Family **Time**: Age, Critical, Current, Date, Duration, Interval **Transportation**: Vehicle, Car, Ship

The Survey provides a very fine-grained categorization of network data. As with the domain context, there is a danger of overspecializing, unless different aspects can in some way be related. It may be important, as well, to consider at least some of the characteristics of data incorporated into the HAT Report taxonomy. For example, many different domains are concerned only with properties that are static or were true at one moment in time, whereas other domains need to consider how those same kinds of properties change over time. The survey might better include among its Basic Information whether the application can highlight changing properties of networks than whether it has been developed in an infection network context or a homeland security context (taking two examples that are not currently listed among the set of domains).

5.6.2 Survey Organization

If the Survey database is to be useful with a Framework-based interface, the categorization of entry terms should be clearly distinguished. The Model-View-Controller approach is helpful in this respect. The Network of interest to a user has properties important for the task. Hence it is important that applications be identified as considering or as ignoring different network properties. The properties are important at the early stage of the framework.

The second stage of the Framework concerns the data. The user may have a dataset that represents the complete network of interest, but usually this is not the case. Parts of the network are incompletely known or unknown. The database should represent whether the application can tolerate uncertainty and incompleteness of information. Likewise, it may be important – or not – that the application continuously updates its analyses as new data arrives, whether sporadically or regularly. As suggested in Table 5-4, many such aspects of data might be valuable to note in the database.

Table 5-4: (HAT Report 3.1) Summary of Data Types

Acquisition	**Streamed**	**Regular**	
		Sporadic	
	Static		
Sources	**Single**		
	Multiple		
Choice	**User-Selected**		
	Externally Imposed		
Identification	**Located**		
	Labelled		
Values	**Analogue**	**Scalar**	
		Vector	
	Categoric (Classic)	**Symbolic**	Linguistic
			Non-Linguistic
		Non-Symbolic	Linguistic
			Non-Linguistic
	Categoric (Fuzzy)	**Symbolic** (Non-Linguistic)	
		Non-Symbolic (Non-Linguistic)	
Interrelations	**User-Structured**		
	Source-Structured		

The third stage of the Framework is Display Requirement. What aspects of the network might the user want to see? This is the link between the raw data and the actual display generation, and it includes all the analytic algorithms that the user might want to deploy. If the user would, for instance, wish to tag each network node with an index of centrality, and each network link with indices of capacity, usage, and availability (three different possible indices of link strength), then the display requirements for an application must include the ability to make those analyses.

The final stage of the Framework is Display Design. When entering information about an application, it should suffice to note what kinds of display designs it supports. Possibly it provides an Application Programming Interface (API) that allows the user to create appropriate displays using separate programs. Possibly the displays are tightly bound to the analyses selected by the user out of the application's repertoire. Possibilities of these kinds, as well as the descriptions of the actual display designs made available by an application, should be incorporated.

All of the above are mere wishes unless the database structure can be well populated with up-to-date information about applications likely to be available to the user. The database therefore should include information about how to acquire the application, as it currently does.

The Survey, even when provided with a Framework-based interface, will be of little use if the information about a high proportion of the applications is skeletal. Filling such a database and keeping it up to date is a lot of work. There are two approaches to this kind of project: contractual and volunteer open-source. An open source project cannot work unless a sufficient community considers the project worth supporting. A "sufficient community" can be either one or two dedicated individuals or a large number of individuals who give some time. A contracted project requires a long-term funding commitment, since keeping the information current is a large part of the work once the initial data have been entered.

5.7 MAPPING THE WORKSHEET TO THE SURVEY

The Framework can be useful without the Survey, as a guide to users as to what to consider when seeking an application for a task. The Survey can be useful without the Framework, if a user wants to know the attributes of an application. The utility of each should multiply many-fold if they can be unified into a working whole, and implemented in functioning software.

Linking the worksheet to the survey will provide a more complete and useful framework. This may best be accomplished by using a mapped scoring function to map between the user's answers within the worksheet and a set of prioritized software applications from the survey – each scored according to most appropriate. This approach is deemed the most useful and realistic as no one application will be a perfect fit for any one user. It is more likely that one application is most appropriate and it is possible that several will fit the bill. The entire framework, once mapped and programmed could be easily posted to a Web-site for user access.

Time did not permit carrying out this mapping within the scope of the current RTG term, but it is recommended that it be carried out by subsequent groups and/or other organizations.

5.8 REFERENCES

[1] NATO IST-063/RWS-010, (2006), "Visualising Network Information", *Workshop Proceedings, RTO-MP-IST-063*, Copenhagen, Denmark, NATO Research and Technology Organisation, 17-20 October 2006.

[2] Bouchard, A. and Vernik, R., (2006), "Characterisation and Showcasing of Network Visualisation Approaches for Command and Control", In *Proceedings of Workshop IST-063/RWS-010 Visualising Network Data*, Copenhagen, Denmark, 17-20 October 2006.

Chapter 6 – PRACTICAL VALIDATION

Practical validation of the framework elements has been carried out by consideration of specific use cases. Four use cases were chosen to cover a variety of topic areas:

- Walsingham (1584): Possible Assassination plot Against Elizabeth.
- Avian Flu on Farms.
- Terrorist Social Network.
- JNDMS – Computer Networks.

For each of these areas, group members acted as users, filling in the worksheet. The results of each now arm the user with the required understanding of their network problem to approach the survey and competently located a visualisation application.

6.1 USE CASES: WORKSHEETS

6.1.1 Walsingham (1584): Possible Assassination Plot Against Elizabeth

The intelligence situation in England of the 1570s and 1580s was remarkably parallel to the post 9/11 situation in the USA. The official reports and other documentation available make possible a remarkably complete record of what was known or could have been known or believed by three sometimes collaborating and sometimes competing intelligence agencies [1]. These agencies were Elizabeth's semi-official intelligence service, run by Sir Francis Walsingham, and two private ones run one by Elizabeth's Privy Councillors (inner Cabinet members) Robert Dudley (the Earl of Leicester), and the other by Dudley's rival, Sir William Cecil (Lord Burghley), succeeded by his son Robert.

The problem for Elizabeth was that her father, Henry VIII, had married Anne Boleyn, Elizabeth's mother, after divorcing his first wife, Catherine of Aragon. In Catholic eyes, this made Elizabeth an illegitimate child with no claim to the throne. To them, several others had a better claim to follow Henry's only legitimate child, Elizabeth's half-sister, "Bloody Mary", who had instigated a reign of terror in trying to return England to Catholicism. Among the possible claimants to the English throne, the most prominent was the Catholic Mary, Queen of Scots. Others were also mentioned, and each was potentially the object of a plot to dethrone or assassinate Elizabeth. However, the Scottish Mary was the major threat, both because she would truly have been the legitimate heir if Elizabeth were deposed (her son did follow Elizabeth on the English throne), and because she had the backing of France, another Catholic country.

Mary had, to put it bluntly, made a disastrous mess of being Queen of Scotland, and had fled to England in the hope of obtaining protection from her English cousin. This had put Elizabeth in a difficult position. She could not allow Mary to be free, because of the strong likelihood of her becoming the nucleus of a serious movement to dethrone Elizabeth. On the other hand, she was unwilling to take strong measures against Mary, initially because Mary had given no offence, but also perhaps from a sense of family obligation and because one ruler should support another in trouble. The compromise was to keep her under house arrest in fairly pleasant circumstances, and to monitor her contacts with the outside world.

Quite apart from the focused threat represented by Mary Queen of Scots, there were real threats from foreign Catholic countries such as France and Spain, who sent agents into England to bolster the Catholic cause and sometimes to foment revolution. Catholics generally were regarded with suspicion following the horrors of the reign of Bloody Mary. Although Elizabeth's official policy was initially one of tolerance, nevertheless Catholics tended to be kept under surveillance, and some independent plots were discovered, both seriously competent and childishly incompetent. Some were supported directly by foreign powers, but most were independent home-grown plots, in much the way that several terrorist plots in recent years have been unrelated to, but inspired by, Al-Qaeda. None came to fruition until the near-success of the 1605 "Gunpowder Plot" two years after Elizabeth's death, the thwarting of which is still celebrated in England 400 years later, every November 5.

The semi-official Catholic subversion was fostered at training schools in France and elsewhere in Europe, and so the English intelligence services had to operate with agents in those countries as well as within England. That there were three competing English intelligence services, all reporting to Elizabeth, meant that each spymaster had to worry whether the agent was possibly a Catholic double agent providing reports of questionable veracity, but also whether he might be working undercover for one of the other agencies and not reporting fully to his official paymaster.

Catholic plotters frequently tried to involve Mary directly, sometimes with her encouragement, sometimes not. Mary's overt correspondence was intercepted, so she resorted to cryptography, steganography and message concealment (for example in supply barrels transported by covert sympathizers). Unknown to Mary, Walsingham's staff were able to decode her messages and to substitute messages of their own using her codes. They were thus able to get early leads on developing plots.

Walsingham had to concern himself not only with Catholic agents who entered the country through normal ports of entry under the guise of ordinary travellers or traders, but also with well-known agents who were smuggled over beaches, as well as with home-grown plotters. At the same time, although all Catholics were under suspicion, and many would have hidden priests and participated in underground religious ceremonies, very few would have plotted actively against Elizabeth.

All the intelligence networks used agents who had been known Catholics and who were turned or subverted by money. Many of these agents worked abroad, often as traders who befriended suspected Catholic agents to develop information about the Catholic social networks. Some attended the Catholic training schools and built up their own networks of contacts. These contacts were sometimes used when a Catholic undercover agent arrived in England to make contact with a cell of potential plotters. The English agent might be imprisoned along with the Catholic, to retain his bona fides in the Catholic social networks. One problem, however, with these turned agents, was that the English authorities could never be sure which side the agent was really working for. Their information had always to be considered unreliable unless corroborated, at least until they had built up a good track record.

Whether at any moment an active plot existed, and who was involved, would have been difficult to determine with the tools available to an Elizabethan spymaster. Walsingham's approach tended toward assuming the worst, with the result that it is quite likely that innocent people were arrested, tried, and in many cases executed. Would the IST-059 Framework have made his task any easier?

Use Case ➔	**Walsingham (1584): Possible Assassination Plot Against Elizabeth**
Worksheet	
Defining Your Problem	
What are you trying to understand? What questions are you trying to answer?	Whether groups of Catholics are developing that might be used by foreign agents in assassination plots against the Queen.
Are you monitoring or influencing a changing situation?	Yes. Both monitoring and influencing by inserting spies into developing networks.
Are you seeking a particular point of information?	Sometimes. Questions arise such as "is person X in contact with person Y".
Are you exploring the network structure for future reference?	Yes. This is the main point, to see changes in the structure that might suggest the development of a plot.
Do you want to be notified when or where a particular condition occurs?	Increases of message traffic and contacts among sub-nets with links to Catholic institutions or known agents.
Does your problem concern the structure of the network or the traffic over the network?	Both, the network being in part defined by message traffic.
Does your problem involve local key points of the network or is it distributed over appreciable sub-nets?	Usually sub-nets associated with key points.
Defining Your Network	
What are the categories of nodes involved?	Persons (blue officials, blue agents, red officials, red agents, and agents of questionable loyalty); places (locations of Catholic institutions, especially training institutions; locations of own major institutions).
For each category of nodes you named, list the relationships or ties that may exist between nodes in that category (not the traffic that passes therein).	Persons are related by message traffic, by institutional history, by affiliation, by family, by employment. Locations are related by persons present, by person traffic, and for teaching institutions by the similarities of concepts taught.
Then, for each pair of categories list the relationships or ties that may exist between pairs of nodes (one from each category).	Person to person messages and face-to-face contact. Person observes relationships among other persons. Meetings occur in locations near significant places.
Does traffic pass between nodes? If so, of what kind (continuous, regular, predictably intermittent, unpredictably intermittent, etc.). If not, what is the nature of the links?	Person to person intermittent unpredictable contact and messages, permanent non-traffic family relationships, alterable non-traffic links of employer-employee and of loyalty ; location to location person traffic, but no traffic on links of conceptual similarity.

Use Case ➔	Walsingham (1584): Possible Assassination Plot Against Elizabeth
Worksheet	
If there is traffic, is the structure of the network defined by the traffic or does it exist independently of whether there actually is traffic over any link?	Message traffic and face-to-face meetings define the network of one class of link. That kind of link does not exist in the absence of traffic. Those links are fuzzy.
For each category of node, does it transform its inputs into different kinds of output. If so, how?	Hard to judge. The possibilities exist.
For each category of node, can the timing of input and output events be related (i.e. are there fixed or variable delays, must two or more inputs occur before an output happens, etc.).	The critical node being actually a sub-net representing a group of conspirators, many contacts and much message traffic must occur before an assassination event is likely.
Important Embedding Fields (Context) of the Network	
What context is important for understanding the network.	Geographical placement and religious affiliation.
Is the most important context a supporting network or a spatially extended area, or something else?	Spatially extended area.
Defining Your Measures	
For each category you named, what about the nodes of that category will you measure?	People will be measured by Loyalty, Affiliation, Religion; Locations will be measured by Affiliation.
For each tie or relationship you named, what about the relationships will you measure?	Frequency of contact, time in a given location.
For any sub-network of your overall network, what about that sub-network will you measure?	Number of people and total loyalty.
For your overall network, what about your overall network will you measure?	Number of people and total loyalty.
Defining Your Resources	
Where will you get your data? (Structured Text/ Databases, Unstructured Text/Documents, Sensor Readings, Other).	Walsingham gets reports from agents of variable trustworthiness about all the other phenomena. Data are incomplete, uncertain, and possibly deliberately false (in the case of reports from double agents).
Are your data predefined for you; can you seek out data to fill gaps in your knowledge; or is the data continually being presented to you in real time?	Data is being presented in real time, and Walsingham can direct agents to seek out particular kinds of data. The networks are derived from the reports of the agents, and are dynamically changeable.

Use Case ➔	Walsingham (1584): Possible Assassination Plot Against Elizabeth
Framework (check each that applies)	
Domain Context	
Task Dynamics and Interactivity	
Real Time	✓
Short Term	
Long Term	✓
Static (One-Shot Analysis)	
Perceptual Modes and Activities	
Explore	✓
Monitor/Control	✓
Search	
Alert	✓
User Role (fill in)	**Advisor to the Queen**
Network Aspects	
Nodes	
Single Mode	
Multimodal	✓
Links	
Single Links	✓
Multiplex	
Metrics	
Single Metric	
Multimetric	✓
Data Characteristics	
Temporal Variance	
Static	
Dynamic	✓
Data Selection	
User-Selected	✓
Interactive	✓
Preset	
Algorithmically Directed	
Data Placement	
Located	✓
Point	✓
Extended	

Use Case ➔	Walsingham (1584): Possible Assassination Plot Against Elizabeth
Framework (check each that applies)	
Labeled	✓
Interactive	
Non-Interactive	
Data Values	
Analogue	✓
Scalar	
Vector	
Categorical	✓
Linguistic	✓
Non-Linguistic	
Data Manipulation	
Interactive	✓
Algorithmic	
Data Interrelations	
User Structured	✓
Algorithmically Structured	

6.1.2 Avian Flu on Farms

In this use case, a health analyst is faced with a possible outbreak of Avian Flu. The user's objective is to track and possibly halt the spread of the disease among poultry in the affected area.

Use Case ➔	Avian Flu on Farms
Worksheet	
Defining Your Problem	
What are you trying to understand? What questions are you trying to answer?	Track and perhaps halt the spread of avian flu among poultry.
Are you monitoring or influencing a changing situation?	Yes.
Are you seeking a particular point of information?	No.
Are you exploring the network structure for future reference?	No.
Do you want to be notified when or where a particular condition occurs?	Yes.

Use Case ➔	**Avian Flu on Farms**
Worksheet	
Does your problem concern the structure of the network or the traffic over the network?	Traffic largely and to a small extent the structure.
Does your problem involve local key points of the network or is it distributed over appreciable sub-nets?	Sub-nets and key points.
Defining Your Network	
What are the categories of nodes involved?	Farms, WildBirds, Barns.
For each category of nodes you named, list the relationships or ties that may exist between nodes in that category (not the traffic that passes therein).	Farms-Farms: Roads, Atmosphere; WildBirds-WildBirds: Proximity; Barns-Barns: Proximity, MemberOfSameFarm.
Then, for each pair of categories list the relationships or ties that may exist between pairs of nodes (one from each category).	Farm-WildBird: FlyOver; Farm-Barn: MemberOf; Barn-WildBird: FlyOver.
Does traffic pass between nodes? If so, of what kind (continuous, regular, predictably intermittent, unpredictably intermittent, etc.). If not, what is the nature of the links?	Trucks, Air and Feces.
If there is traffic, is the structure of the network defined by the traffic or does it exist independently of whether there actually is traffic over any link?	The whole problem is about the viral traffic – that is the basis of the network.
For each category of node, does it transform its inputs into different kinds of output. If so, how?	No.
For each category of node, can the timing of input and output events be related (i.e. are there fixed or variable delays, must two or more inputs occur before an output happens, etc.).	There is some small amount of time between input of the virus and output of the virus.
Important Embedding Fields (Context) of the Network	
What context is important for understanding the network?	The geographical locations of the farms and the vehicle/bird traffic between on which the virus rides.
Is the most important context a supporting network or a spatially extended area, or something else?	A supporting network.
Defining Your Measures	
For each category you named, what about the nodes of that category will you measure?	Farm: InfectedBirds (Y/N); Barn: InfectedBirds (Y/N); WildBirds: Infected (Y/N).

Use Case ➔	Avian Flu on Farms
Worksheet	
For each tie or relationship you named, what about the relationships will you measure?	Roads: TrafficIntensity, BirdSales; Atmosphere: Distance, WindDirection; Proximity: Distance; MemberOfSameFarm: Y/N; Flyover: Probability; MemberOf: Y/N.
For any sub-network of your overall network, what about that sub-network will you measure?	% of infected barns in a region; Marketing convergence within a region.
For your overall network, what about your overall network will you measure?	VirusFree (Y/N); VirusTypePresent.
Defining Your Resources	
Where will you get your data? (Structured Text/ Databases, Unstructured Text/Documents, Sensor Readings, Other).	Sensor readings, documents, global public health information network, dept of agriculture.
Are your data predefined for you; can you seek out data to fill gaps in your knowledge; or is the data continually being presented to you in real time?	We can seek out data to fill in the gaps but the data is always coming in.
Framework (check each that applies)	
Domain Context	
Task Dynamics and Interactivity	
Real Time	✓
Short Term	✓
Long Term	
Static (One-Shot Analysis)	
Perceptual Modes and Activities	
Explore	
Monitor/Control	✓
Search	
Alert	✓
User Role (fill in)	**Health Analyst**
Network Aspects	
Nodes	
Single Mode	
Multimodal	✓
Links	
Single Links	
Multiplex	✓

Use Case ➔	Avian Flu on Farms
Framework (check each that applies)	
Metrics	
Single metric	
Multimetric	✓
Data Characteristics	
Temporal Variance	
Static	
Dynamic	✓
Data Selection	
User-Selected	✓
Interactive	✓
Preset	
Algorithmically Directed	
Data Placement	
Located	✓
Point	✓
Extended	✓
Labeled	✓
Interactive	
Non-Interactive	
Data Values	
Analogue	
Scalar	✓
Vector	✓
Categorical	✓
Linguistic	✓
Non-Linguistic	
Data Manipulation	
Interactive	✓
Algorithmic	
Data Interrelations	
User Structured	
Algorithmically Structured	✓

6.1.3 Terrorist Social Network

In this use case, an intelligence analyst has the task of tracking and understanding a particular terrorist cell that may be operating in their area of responsibility. In particular, the analyst is seeking to locate who within the cell is "in charge" and who may be financing the cell. For this scenario, the structure of the group is more important than the traffic passing between them.

Use Case ➔	Terrorist Social Network
Worksheet	
Defining Your Problem	
What are you trying to understand? What questions are you trying to answer?	I need to understand the structure of a particular terrorist cell – who the players are, how is in control, etc.
Are you monitoring or influencing a changing situation?	Yes.
Are you seeking a particular point of information?	Yes – who is in charge and who is the financier.
Are you exploring the network structure for future reference?	Yes.
Do you want to be notified when or where a particular condition occurs?	Yes.
Does your problem concern the structure of the network or the traffic over the network?	The structure more than the traffic.
Does your problem involve local key points of the network or is it distributed over appreciable sub-nets?	Key points – particular people.
Defining Your Network	
What are the categories of nodes involved?	People, organizations.
For each category of nodes you named, list the relationships or ties that may exist between nodes in that category (not the traffic that passes therein).	People-People relationships are: SubordinateTo, RelatedTo, LocatedWith; Organization-Organization relationships are: LinkedTo, SubordinateTo, FunderOf.
Then, for each pair of categories list the relationships or ties that may exist between pairs of nodes (one from each category).	People-Organization relationships are: MemberOf, LeaderOf, FunderOf.
Does traffic pass between nodes? If so, of what kind (continuous, regular, predictably intermittent, unpredictably intermittent, etc.). If not, what is the nature of the links?	Traffic passing between nodes includes: Money and Information and is unpredictable.
If there is traffic, is the structure of the network defined by the traffic or does it exist independently of whether there actually is traffic over any link?	Independent of any traffic.

Use Case ➔	**Terrorist Social Network**
Worksheet	
For each category of node, does it transform its inputs into different kinds of output. If so, how?	No.
For each category of node, can the timing of input and output events be related (i.e. are there fixed or variable delays, must two or more inputs occur before an output happens, etc.).	No.
Important Embedding Fields (Context) of the Network	
What context is important for understanding the network	Geographical locations.
Is the most important context a supporting network or a spatially extended area, or something else?	Spatially extended.
Defining Your Measures	
For each category you named, what about the nodes of that category will you measure?	People will be measured by: certainty of existence. Organizations will be measured by: number of members and total funding.
For each tie or relationship you named, what about the relationships will you measure?	Most will be binary (T/F) but funding will be measured in USD.
For any sub-network of your overall network, what about that sub-network will you measure?	Sub-networks will be measured in size and total funding like an organization.
For your overall network, what about your overall network will you measure?	The total network will be measured in size and total funding like an organization.
Defining Your Resources	
Where will you get your data? (Structured Text/ Databases, Unstructured Text/Documents, Sensor Readings, Other).	Unstructured text from disparate sources.
Are your data predefined for you; can you seek out data to fill gaps in your knowledge; or is the data continually being presented to you in real time?	We can seek out data to fill in the gaps but there will be a lot of missing data. Data is always coming in, but not in predictable increments.
Framework (check each that applies)	
Domain Context	
Task Dynamics and Interactivity	
Real Time	✓
Short Term	✓
Long Term	✓
Static (One-Shot Analysis)	

Use Case ➔	Terrorist Social Network
Framework (check each that applies)	
Perceptual modes and activities	
Explore	✓
Monitor/Control	✓
Search	
Alert	✓
User Role (fill in)	**Analyst**
Network Aspects	
Nodes	
Single Mode	
Multimodal	✓
Links	
Single Links	
Multiplex	✓
Metrics	
Single Metric	
Multimetric	✓
Data Characteristics	
Temporal Variance	
Static	
Dynamic	✓
Data Selection	
User-Selected	✓
Interactive	✓
Preset	
Algorithmically Directed	
Data Placement	
Located	✓
Point	✓
Extended	✓
Labeled	✓
Interactive	✓
Non-Interactive	
Data Values	
Analogue	
Scalar	✓

Use Case ➔	**Terrorist Social Network**
Framework (check each that applies)	
Vector	✓
Categorical	✓
Linguistic	✓
Non-Linguistic	
Data Manipulation	
Interactive	✓
Algorithmic	✓
Data Interrelations	
User Structured	✓
Algorithmically Structured	

6.1.4 JNDMS – Computer Networks

This use case is focused on maintenance and sustainment of a given computer network. The network may not necessarily be under attack, but may experience "events" that reduce or otherwise compromise network operation, availability of services, connectivity, etc.

Use Case ➔	**JNDMS – Computer Networks**
Worksheet	
Defining Your Problem	
What are you trying to understand? What questions are you trying to answer?	The impact of network events on operations; The event is loss of a host server.
Are you monitoring or influencing a changing situation?	Monitoring.
Are you seeking a particular point of information?	Yes – when and at what node an event has occurred.
Are you exploring the network structure for future reference?	No.
Do you want to be notified when or where a particular condition occurs?	Yes.
Does your problem concern the structure of the network or the traffic over the network?	Traffic – and minimally structure in the sense of event isolation.
Does your problem involve local key points of the network or is it distributed over appreciable sub-nets?	Key points but leading to affects on sub-nets.
Defining Your Network	
What are the categories of nodes involved?	Devices, Client Software, Server Software, Operations.

Use Case ➔	JNDMS – Computer Networks
Worksheet	
For each category of nodes you named, list the relationships or ties that may exist between nodes in that category (not the traffic that passes therein).	Device-to-Device: physical link, information link; ClientSoftware-to-ClientSoftware: none; ServerSoftware-to-ServerSoftware: none.
Then, for each pair of categories list the relationships or ties that may exist between pairs of nodes (one from each category).	Device-to-ClientSoftware: ResidesOn; Device-to-ServerSoftware: ResidesOn; ClientSoftware-to-ServerSoftware: Requires.
Does traffic pass between nodes? If so, of what kind (continuous, regular, predictably intermittent, unpredictably intermittent, etc.). If not, what is the nature of the links?	Yes – regular data traffic.
If there is traffic, is the structure of the network defined by the traffic or does it exist independently of whether there actually is traffic over any link?	Structure is independent of the traffic but exists to allow the traffic.
For each category of node, does it transform its inputs into different kinds of output. If so, how?	No.
For each category of node, can the timing of input and output events be related (e.g. are there fixed or variable delays, must two or more inputs occur before an output happens).	No.
Important Embedding Fields (Context) of the Network	
What context is important for understanding the network.	The hardware base of the system and its capabilities/ vulnerabilities.
Is the most important context a supporting network or a spatially extended area, or something else?	Supporting.
Defining Your Measures	
For each category you named, what about the nodes of that category will you measure?	Devices: status (up/down/degraded); ClientSoftware: ServerResponse; ServerSoftware: ServerResponse.
For each tie or relationship you named, what about the relationships will you measure?	PhysicalLink: BandwidthConsumption; InformationLink: Bidirectionality; ResidesOn: yes/no; Requires: yes/no.
For any sub-network of your overall network, what about that sub-network will you measure?	Status (up/down/degraded).
For your overall network, what about your overall network will you measure?	Status (up/down/degraded).

Use Case ➔	JNDMS – Computer Networks
Worksheet	
Defining Your Resources	
Where will you get your data? (Structured Text/ Databases, Unstructured Text/Documents, Sensor Readings, Other)	Structured data from system diagnostics.
Are your data predefined for you; can you seek out data to fill gaps in your knowledge; or is the data continually being presented to you in real time?	Predefined but presented in real time.
Framework (check each that applies)	
Domain Context	
Task Dynamics and Interactivity	
Real Time	✓
Short Term	✓
Long Term	
Static (One-Shot Analysis)	
Perceptual Modes and Activities	
Explore	
Monitor/Control	✓
Search	✓
Alert	✓
User Role (fill in)	
Network Aspects	
Nodes	
Single Mode	
Multimodal	✓
Links	
Single Links	
Multiplex	✓
Metrics	
Single Metric	
Multimetric	✓
Data Characteristics	
Temporal Variance	
Static	
Dynamic	✓

Use Case ➔	JNDMS – Computer Networks
Framework (check each that applies)	
Data Selection	
User-Selected	✓
Interactive	
Preset	✓
Algorithmically Directed	
Data Placement	
Located	✓
Point	✓
Extended	
Labeled	✓
Interactive	
Non-Interactive	✓
Data Values	
Analogue	✓
Scalar	✓
Vector	
Categorical	✓
Linguistic	✓
Non-Linguistic	
Data Manipulation	
Interactive	
Algorithmic	✓
Data Interrelations	
User Structured	
Algorithmically Structured	✓

6.2 COMMENTS AND RECOMMENDATIONS

After filling in the table, it seemed that the answers did not adequately express the complexity of the task. There seemed to be several different tasks that should have been addressed more extensively. It is probable that a software-based question-answer interface would support this better than a paper worksheet that demands small boxes to be filled in. Nevertheless, some of the issues were clarified by the process of answering the questions and for which different answers in one part of the worksheet are not connected with the sub-answers in another part. **The exercise was very useful in helping the overwhelmed users better understand their cases, but could be still more effective if this worksheet were to be instantiated as a Web-based software tool.**

6.3 REFERENCES

[1] Haynes, A., "The Elizabethan Secret Services", Sutton Publishing, 1992/2004.

Chapter 7 – ASSOCIATED ACTIVITIES: WORKSHOPS AND THE NETWORK OF EXPERTS

During its mandate, IST-059/RTG-025 sponsored one NATO workshop, held in Copenhagen, Denmark in October 2006. The Programme Committee included members of the RTG and the work was considered to be integral to that of the RTG. In addition, the Visualisation Network of Experts (Vis N/X) conducted a workshop in each of 2005, 2007 and 2008, years in which the RTG did not sponsor any NATO workshops. This followed the pattern used by the predecessor RTGs during which sponsored NATO workshops were interspersed with Vis N/X workshops on an annual basis.

7.1 METHODS, PARTICIPATION AND RESULTS

Each workshop, in addition to holding general plenary presentations and discussions of research and applications, used small "syndicates" or "working groups" to examine specified topics. The selected topics had in many cases been introduced by syndicates at an earlier workshop, and were considered by the RTG to be important to develop further. For example, the Vis N/X Workshop in 2003 was asked to further develop counter-terrorism ideas reported by one of the syndicates at the Halden NATO workshop in 2002.

So far as feasible, each syndicate included representatives of the scientific research community, the developer community, and the military user community. The intention was to ensure that the ideas developed in the intensive work of the syndicate would be militarily relevant, scientifically defensible, and developmentally feasible. In most cases, this attempt was reasonably successful, though the military user community was, as might be expected, less well represented at the Vis N/X workshops than at the NATO workshops.

Many of the same researchers and several of the more senior military attended more than one workshop, which made for easier communication at the later ones, as the researchers became better aware of military needs, and the military of scientific possibilities.

Several of the same issues emerged at each workshop. One of the critical ones was the problem of displaying relationships. Almost all action requires the actor to assess relationships. Even in such a trivial everyday action as to transport an object, the relationship between the object's size and the capacity of containers, the relationship between the terrain and the available transport mechanisms, the relationship between the cost of replacing the object and the risk of loss, are only a few of the relationships that must be implicitly or explicitly considered. Furthermore, there are second-order relationships among even these simple ones. Techniques for displaying relationships in a manner that aids visualisation are not well developed.

The concept of relationships inevitably implies the concept of networks. If A, B, and C each has some kind of relationship with the others, the set of relationships forms a network in which A, B, and C are "nodes" and the relationships among them are "links". Some networks have well-defined link structure, in the sense that if A is linked to B and not to C, an action by A is propagated to B and not to C. Other networks have a less well-defined structure. For example, A may act on an environment that is available for inspection by B and C, but which will not necessarily be so inspected until much later, if ever. Yet again, some networks may be linked by momentary broadcast; A may broadcast a message, but B and C will receive it only if their receivers are turned on at that specific moment. In neither of these latter cases is the network structure defined beyond a probabilistic statement about the existence of the link.

Networks have emergent properties beyond those implied by the nature of the nodes and of the relationships among the nodes. For example, the nodes in a network may be connected randomly, they may be connected as a set of branches radiating from a single hub, they might be connected as a hierarchy of networks connected locally in random way but with the local nets connected through higher-order networks of hubs, or they might have other statistically describable patterns of linkage. These patterns have strong effects on the ways networks behave, and affect the requirements for their display. For example, a road network includes roads with widely different traffic-carrying capacities, from multilane expressways to rutted cart tracks. A display of the road network for one purpose such as showing the fastest routes between major cities might include only the expressways, whereas a display for hikers might show the cart tracks and indicate the expressways only as obstructions.

When the different properties of the nodes and links in a network are added to the structural differences among networks, the display requirements become very challenging. At the 2004 Toronto Workshop, one of the working groups attempted to draft a set of abstractions of node and link properties that could be used in developing display requirements for different purposes. They then suggested which of these properties would probably be important for user purposes in different applications. As examples, they chose the application areas of counter-terrorism, information assurance, and logistical analysis.

Networks not only have static properties defined by the analysis of their node-link structure and the natures of the nodes and links, they also have dynamic properties. Events in one part of a network may propagate along the links to other parts of the network. Their influence may dissipate over time, may grow and then dissipate, may cause oscillations, or may develop chaotically. Visualisation requirements may include the provision of mechanisms for users to assess the probable future evolution of the effects of different action choices. One obvious area in which this would be important is in the delicate socio-political networks involved in peacekeeping operations.

An earlier 2004 Toronto workshop (IST-043) had as its theme “The Common Operational Picture”. It became obvious from the work of several of the working groups that the concept of a common operational picture left much to be desired. People collaborating in any venture, whether in a battlefield that contains a definable enemy, or in the design of a complex device, need to know what the other actors know and intend, but they do not need to see the same “picture”. Each collaborator has a different purpose. To fulfil that purpose, s/he needs to know the overall objective and where the purpose fits into that objective, and needs to know where the purposes of others also fit, especially if their actions influence the environment in which she would act. The terms “Common Operational Environment” and “Common Knowledge of Intent” might be more appropriate than “Common Operational Picture”.

At every workshop, the importance of real-time interaction between the user and the display was emphasised. The ability of the user to control not only the display content, but also the manner of display (as, for example, the viewpoint in a simulated scene), greatly affects the effectiveness of the user’s visualisation. At a very early Vis N/X workshop, for example, the claim was made that in a display of the message-passing and inheritance structure of a moderately large software system, a passively rotated 3-D display allowed the user to visualise about twice the amount that could be visualised from a 2-D display, whereas an interactive 3-D display in which the user cold manipulate the viewpoint increased the advantage to a factor of between five and seven.

As the several workshops pointed out, not only the presentation on the display surface, but also the user’s interaction with the display is likely to be critical to the success of a visualisation system. For this reason, IST-021/RTG-007, following the lead of its predecessor groups, used a reference model based around the concept of a layered structure of interaction. This “VisTG Reference Model” presented a framework within which different display technologies could be deployed, but that also suggested to designers and evaluators the possibilities for displaying not only the data to be visualised, but also the control the user should be given over the displays in support of different kinds of purpose.

7.2 NATO WORKSHOP COPENHAGEN: IST-063/RWS-010 WORKSHOP ON VISUALISING NETWORK INFORMATION

The IST-063/RWS-010 workshop on Visualising Network Information [1] took place in Copenhagen in October 2006. The term “network” includes both “physical” networks – e.g. information and service infrastructural networks – as well as “conceptual” networks – e.g. social networks which show the interactions and organizational relationships among their elements.

The workshop intended to bring together those who use network analysis system, those who develop them, and those who make the systems more usable and effective. A core objective was to have users talk with developers and researchers. The workshop was intended as a forum for commanders and staff officers to describe the pros and cons of current systems supporting network visualisation, which should help guide future military visualisation and research and development. The aim was to be multidisciplinary since both human factors and technological innovation collaborate in improving visualisation systems. The workshop intended to identify problems to which there are as yet no solutions, but where solutions seem possible.

The workshop satisfied the objectives to a certain degree. It had 28 invited active participants of the following mix:

- 4 military;
- 7 government;
- 8 academia; and
- 9 industry.

Some of the participants fit into more than one category.

The participants came from the following countries:

- 9 USA;
- 8 CAN;
- 4 DNK;
- 3 GER;
- 2 NOR;
- 1 GBR; and
- 1 SWE.

The workshop was organised in sessions consisting of presentation of a number of papers, questions to the papers, and plenary discussion of the topic of the session. In addition over 30% of the workshop was devoted to focused working groups.

The first session after the introduction was the 2 keynotes by:

- Professor Kathleen Carley from Carnegie Mellon University, USA.
    - Title: A Dynamic Network Approach to the Assessment of Terrorist Groups and the Impact of Alternative Courses of Action.
- Colonel (Retired) Randy Alward from Canada.
    - Title: A Need for Better Network Visualisation.

The following sessions had the following topic areas:

- General/theory (2 sessions / 5 papers);
- Security/Defence (3 sessions / 8 papers); and
- Medical (1 session / 3 papers).

The work groups worked in parallel sessions on the following topics:

- Framework Survey Integration;
- Reliability and uncertainty in situation awareness of Network Visualisation (3 groups); and
- Vulnerability and network Analysis.

On the last day the working groups reported back to the workshop in a plenary session.

The following points are particularly significant.

Social Network Analysis (SNA) is the mapping and measuring of relationships and flow between people, groups, organizations, computers, Web-sites, and other information/knowledge processing entities. The nodes in the network are the people and groups while the links show relationships or flow between the nodes. This makes SNA an important tool when fighting terrorism. SNA is multidisciplinary involving several aspects of both information science and human factors. The first keynote speaker Prof. Kathleen M. Carley from Carnegie Mellon University has taken SNA a step further to Dynamic Network Analysis (DNA) where the relationships are dynamic. It is still an emergent technology, but one to watch.

Network Centric Warfare, Network Enabled Operations, Network Enabled Capability, Network Centric Operations are the terms different nations use for the same concept, transforming their procedures to take advantage of the Information Age. It is about using networks to speed up and improve the C2 process. It is necessary to maintain the networks in support of operations, and to do that C2 must be extended to networks and cyberspace. An integrated information environment is required for this, and usually there are four environments:

- The unclassified Internet;
- The designated Intranet;
- The classified Command and control systems; and
- The special classified networks for areas such as Intelligence.

Today network visualisation shows a logical view of the networks with an indication that data is or is not flowing between routers (green meaning data is flowing and red meaning it is not). It is not an acceptable indication of availability. There might not be adequate bandwidth to pass the traffic. The network operators are happy as long as the indications are green, when in fact the user's needs might not be met.

There is no geographical representation of the network, so the network operations staffs have no idea as to whether an outage affects an operation or not. How can they prioritize which problem to address first?

They are, in general, unaware that upgrades are being rolled out, which are often the cause of problems.

For all intents and purposes our Network staffs are blind, unable to make timely, prioritized decisions regards network repairs. This is certainly an unacceptable condition for Network Enabled Operations!

Network technology allows the commander to centralize command and control but this goes against the concept of decentralized elements that carry out operations. It is doctrine, not capability, that will keep the joint chiefs out of the commander's backyard.

Visualising networks with thousands of nodes will overload the user, so it is necessary to reduce the complexity of the network. It is even difficult to visualise networks of more than 50 nodes. Several ways to attack this problem were discussed, but no solution was advanced.

The Medical session showed that co-operation on network visualisation between the medical science and information technology communities will benefit both in several areas:

1) There is similarity between computer viral transmission and biological viral transmissions. They both attempt to track and investigate the infection after the infection has begun. If the virus is spread via email then the parallels are strong, but if it is a specifically targeted attack (by proxy or otherwise) then the parallels may not be so strong. Other parallels include susceptibility or lack of susceptibility based on inoculation and/or temperament of the user. But there is a specific difference between computer and biological viruses in that some computer viruses are programmed to strike on specific days. There are parallels to bio-terrorism because computer viruses are human-induced and a global infection can be very quick. The comparison is reasonable because the computer model can take in parameters for susceptibility. However, proximity is not a factor for computer virus propagation, so an analogy to a neural network would fit more appropriately.
2) There are strong parallels between Information Assurance practices and the discovery and containment of sexually transmitted diseases and infections.
3) There could be value in a real time social network display during an outbreak so that actions can be taken to contain it. Visualisations can help to identify nodes with different characteristics.

Participants reported that it was a productive workshop, but there was a disappointingly small number of participants from Europe as compared to previous workshops in this series (both NATO and N/X). In spite of this there was a stimulating interchange throughout the workshop.

7.2.1 Conclusions

Most network information sets do not take "time" into account.

Network analysis and visualisation are:

1) Important tools in the fight against terrorism
2) Useful for tracking disease and attacks on computer systems (virus, etc.).

Better view of the networks and what is occurring on them is needed. Multiple views of the network are needed:

- A logical view that shows communications links, routers, servers, firewalls and applications.
- A physical view that overlays the logical view on a geographical representation.
- A transactional view that shows if the various applications are functioning. Is logistics delivering just in time supplies? Are invoices being paid on time?
- An operational view that shows commanders and staff are able to use the networks to gain the advantage that Network Enabled Ops promises. Network staff can prioritize restoral on the basis of operational priorities.

There is not enough collaboration between the academic researchers and the defence community. The academic researchers would like realistic information to test their systems. If the systems are not tested on realistic information they might not be developed into useful systems for defence.

There is a critical need for advancing network Command and Control.

Progress has been marginal in the Network Operations Centres.

The two topics chosen by the workgroups “Reliability and uncertainty in situation awareness of Network Visualisation” and “Vulnerability and Risk Assessment” are important topics needing more research.

7.2.2 Recommendations

Contact between users, developers and researchers should be encouraged.

The problems need to be defined in ways that allow even civilian researchers to work on them.
Better ways to get laundered data to researchers/modellers should be established.

Operational studies and analyses of visualisation needs from the analyst’s and commander’s viewpoints should be initiated. A seamless environment across Net Ops Centres and R&D Labs should be created and R&D results should be used in the Ops Centres as soon as possible.

One “low hanging fruit” is to use higher resolution display technology or an appropriately matrixed array of high resolution displays oriented in such as manner as to maximize the information availability without “overloading” the user. Further research may improve how displays of this calibre can be organized to maximize the information output without creating an overload situation.

Experimentation with 3-D visualisation should be encouraged.

More research on “Reliability and uncertainty in situation awareness of Network Visualisation” and “Vulnerability and Risk Assessment” should be encouraged.

Concerning uncertainty and reliability the following is necessary:

- A clear definition of reliability and uncertainty is needed.
- Development of visualisation concepts and prototypes, defining what uncertainty and reliability conveys.

- Conduct experiments with representations of uncertainty and reliability.
- Development of consistent techniques for determining uncertainty and reliability.
- Development of intuitive techniques for visualising uncertainty and reliability.

7.2.3 Post-Workshop Review

A post-workshop review was held. Here are some of the salient comments:

- The role of the formal papers was questioned. It was suggested that since the objective of the workshop was to develop ideas through discussions and working groups, the main reason for having the formal papers was to evaluate the appropriateness of issuing an invitation to the participant, not to provide a publication vehicle for work done outside the workshop.
- The discussion periods might have been too short. In other workshops it has been noted that longer discussions tend to develop ideas late in the discussion periods. In the present workshop, the discussion periods seldom got beyond questions addressed to the talkers of the previous session.
- Sessions were sometimes rushed, with speakers overstepping their allotted time, often by large factors. Something must be done to keep the presentation times short, so as not to cut into the working sessions.
- No reason was found for the relatively small number of European participants, though it was noted that the Call for Participation had not been effectively distributed in some nations. Most attendees found out about the Workshop either by having been members of the N/X or from word of mouth.
- The decision to pre-assign participants to working groups which selected their own topics from a menu turned out well and should be repeated.
- It was suggested that if key participants could be “locked in” early, their participation could be noted in a second pass call for participation.
- At early workshops in the series, it was possible for a participant to request a short presentation period at the start of a (longer) plenary discussion session, if they thought that the provocation presentations raised points to which their own work provided relevant commentary that would aid the following discussion.

7.3 NETWORK OF EXPERTS WORKSHOPS AND ACTIVITIES

The Visualisation Network of Experts (Vis N/X or N/X) was created in the mid-1990s as an informal technical advisory group for a predecessor visualisation research group to IST-059. The Vis N/X supported each subsequent visualisation RTG and most recently supported IST-059 in realizing its mandate. The Vis N/X has held nine workshops in conjunction with meetings of the “parent” RTG, with the aim of focusing needed research in this area.

The N/X holds workshops on Visualisation in years that VisTG does not sponsor official NATO Workshops. Typically, the N/X Workshops are held alternately in North America and in Europe. N/X Workshops usually last two days, and are held in conjunction with a VisTG meeting. The workshops include breakout sessions in which small groups work intensively on a topic related to the meeting theme. These sessions often result in recommendations to VisTG.

The Vis N/X has a select membership. It includes the members of the current patron RTG plus invited experts. Its original members were identified at the Brussels workshop of 1994 and were subsequently invited to form the Vis N/X by RSG-30 when it was created in 1996. The Vis N/X expands by inviting other experts identified by its members: any Vis N/X member may recommend an expert for membership provided that that expert comes from a NATO or PfP country. At the end of December 2008, there were 95 members of the Network of Experts, not including the members of the RTG.

At its inception, the Vis N/X realized a new concept in NATO research discussions and activities. In operation it offers an unofficial forum for researchers to exchange information, data and expertise. It carries some of the advantages that the NATO umbrella can offer, while avoiding some of the problems with more formal arrangements, including some Governments' occasional reluctance over the last decade to join in official arrangements.

7.3.1 N/X Bonn

The N/X held a meeting at FGAN-FKIE, Wachtberg-Werthhoven, DEU during October, 2005. The topic was "Social Network Analysis and Visualisation for Public Safety". Social network analysis provides a means to study the varied and diverse interactive relationships among individuals, organisations, groups and countries. The impact of the social network whether on the individual or on the nation plays an important role influencing individual or collective physical, environmental and public safety. Social network analysis is used in public safety, for instance, detecting and tracking terrorists or extremists for CBRN (chemical, biological, radiological and nuclear) threats, domestically generated threats, organised crime or civil emergency. Researchers working on social network analysis have also, for example, identified how infectious diseases such as SARS or STDs can be spread among individuals across different social groups and communities.

The participants of the workshop addressed the above from different perspectives and using different approaches but collectively they complemented each other. Notable among them:

- A new approach for providing information for intelligence analysts, military commanders, individual solders and others using the concept of custom ontologies based on each users query was discussed: a concise and organized knowledge set with the appropriate visualisation tool to assist exploration and facilitate intelligence assimilation is provided to users, c.f. convention approach of searching thousands relevant and irrelevant documents or building a large all-encompassing ontology.
- Other presenters suggested a suite of several means to resolve knowledge creation problems from intelligence and other data: increase the bandwidth, develop a cyber equivalent of a fly-through data approach, and conserve analyst attention, focus on the negative space and adapt to individual users.
- Work at FGAN focused on the identification and visualisation of relations extracted from ISR-messages (Intelligence, Surveillance and Reconnaissance) using the FGAN's J2-Database (DBEins) for message acquisition and xGErD, a network representation of relations.
- Health Canada developed tools including VITA and the Master Battle Planner [developed by QinetiQ] to manage and visualise network relationship and logistic analysis for infectious disease outbreak management, which are applicable for the management and analysis of CBRN and other public threats.

Unlike previous Vis N/X workshops, this one included an appreciation, taking as an example the technical review done normally in more formal NATO Workshops. The appreciation by David Zeltzer [USA] and Marcus Lem [CAN] highlighted research matters arising; they further commented on both what had gone well and what could be improved with the format. The lessons learned were useful for the subsequent NATO workshop in Copenhagen DNK and the Vis N/X workshops in El Segundo USA and Malvern GBR.

The agenda for this workshop and links to presentations and abstracts can be found in Annex K.

7.3.2 N/X El Segundo

The Vis N/X held a workshop at Aerospace Corp, El Segundo, in November, 2007 on "Network Analysis for Simulation and Prediction". At this workshop there was a series of keynote addresses, provocations, breakout working groups and plenary sessions over three days, ending with a site visit to Jet Propulsion Laboratory. Throughout, there were informal technical discussions over meals and breaks.

Amy K.C.S. Vanderbilt, USA, set the outstanding questions in the plenary:

1) How do we usefully assess if various types of networks' are predictable over time?
2) When and how much are the various networks predictable?
3) Can network prediction tools and algorithms be sufficiently tested within simulated or modeled networks? How certain can we be of the results of such models?
4) What role does visualisation play in measuring and understanding network predictability (or lack thereof) and the predictions (if any)?

"It comes down to this," she continued:

> There are efforts from every branch of the military seeking to predict the behavior of terrorist cells and other networks given various influence factors. However, I am not so sure that all such networks behave with any degree of predictability. I am also not sure that they don't. This is a question that has not been addressed – perhaps because a viable definition/measure of predictability has not been formulated. But at the same time it is a question that needs to be addressed.

Workshop organizers especially invited discussions of visualisation of graph/network simulation and prediction that span disciplines and afford useful applications in new domains such as the application of centrality measures in social networks to vulnerability assessment in computer networking environments or detection of key nodes in communications networks employed by adversaries.

The workshop format was similar to earlier Vis N/X and NATO workshops held at Penn State University USA, Halden NOR, Toronto CAN, Wachtberg-Werthhoven DEU and Copenhagen DNK. That format makes extensive use of "provocations" – presentations lasting no more than 15 minutes intended to stir discussion; for this Workshop we asked that a provocation pose a question or main point you consider important, for discussion. Discussion followed immediately for 10 minutes, extended later over coffee or meals.

As well as the provocation format, the workshop committee affected short position papers were accepted from those who prefer that format. The plenary decks and papers are included in the Proceedings.

Breakout groups addressed topics as follows:

- Developing a framework for network visualisation, to accommodate various ways of treating and understanding static and dynamic networks;
- Developing network datasets for research and understanding; and
- Visualising uncertainty in network contexts.

Decks and presentations form the breakout groups are included in Proceedings as well.

Not surprisingly, the workshops did not answer the all of the questions asked in the plenary but there was a broad feeling that real progress was made, particularly in defining the needs for datasets and recognizing the roles for modeling network dynamics and visualising the results to attack the problem suite.

The agenda for this workshop and links to presentations and abstracts can be found in Annex L.

7.3.3 N/X Malvern

The format [provocations followed by discussion, accompanied by related breakout groups] had been used several times several times previously but it seems to have proved its worth once again; as a result, no change in format is planned. The agenda is attached, and the provocation decks will be made available on the Web and in proceedings as soon as possible. Likewise attached are decks from breakout groups' presentations.

Breakout groups addressed four topics selected by popular vote from a dozen topics suggested by the Workshop Committee. Topics addressed were:

- Experimental design to evaluate specific visualisations' utility;
- Multimodal and multirelation networks;
- Representing uncertainty in network visualisation [this was decomposed into representing uncertainty about the topology and structure of the network and representing uncertainty about the individual node properties [including capacity] and individual link properties [including traffic density and direction]; and
- Use of information theory to describe and quantify visualisations of networks.

Decks are posted and well worth reading. In particular, Group 4 above feels that they have started a concerted effort to create a "unified theory of networks" capable to model networks as disparate as railway lines and brains.

IST-059 met briefly after the meeting and generated several comments and recommendations.

- This was the usual plethora of random philosophical thought, with the usual excellent results, generating new ideas in the field.
- I love to meet with new people with new ideas and generate good new ideas during these workshops!
- There's a momentum developed here for practical near term deliverables.
- Longer meetings would be better to develop the thoughts and discussions from all the provocations.
- The workshop pushed me to apply these principles to my work in a way that I have not been pushed before.
- A birds-of-a-feather group where things really get done, unlike those at larger conferences.
- I will participate in an annual visual evaluation exercise next week. And I will use [the RTG] framework there.
- My colleagues have run into a dead end in these areas for some years. I will insist that they employ the ideas we have generated here.
- Great impact on my research.

- As ever, I enjoyed it. We think we've pushed the state of the art this time, and we do that every time.
- One of the most profitable academic exercises I ever have become involved with.
- In spite of the high quality of the provocations, there should have been fewer of them.
- A pre-publication of the abstracts would be useful.
- It went very well – external feedback is also positive. Participants felt it was a good experience and were glad to have taken the time to come here. People want to get more actively involved with us.

The agenda for this workshop and links to presentations and abstracts can be found in Annex M.

7.4 WEB-SITES AND MAILING LISTS

Continuing the precedent set by IST-013 and IST-021, IST-059 operated a public and a private Web-site of RTG information as well as a private Web-site that supported collaboration among the members of the group. This latter site included a private Forum for technical discussion, as well as a Wiki for more permanent material. Collaboration through these media enabled development of several novel concepts leading toward the development of the "Unified Theory of Networks" that may ultimately be an outcome of the Framework development.

Both IST-059 and the Vis N/X operated mailing lists which were only accessible to members of the groups. A considerable volume of the work of the RTG was done or promoted through these various modes of electronic interaction.

7.5 REFERENCES

[1] NATO IST-063/RWS-010, (2006), "Visualising Network Information", *Workshop Proceedings, RTO-MP-IST-063*, Copenhagen, Denmark, NATO Research and Technology Organisation, 17-20 October 2006.

Chapter 8 – IMPACT AND VALUE ADDED

8.1 INTERNATIONAL COLLABORATION AND INFLUENCE

- **NATO – TTCP**: Martin Taylor attended a meeting of TTCP C3I TP2 Command Visualisation. A presentation was given on IST-059, and the discussion seemed to lead to the idea that there might be some benefit in continued interaction.
- **Norway – Sweden**: Thomas Porathe, Sweden, participated in the NATO workshop in Copenhagen, spurring contact with delegate Jan Terje Bjørke of Norway for collaborative research exchange.
- **Norway – Canada**: Dragos Calitoiu, Canada, through a Vis N/X workshop, contacted Jan Terje Bjørke of Norway for collaborative research exchange. Activity is ongoing.
- **Canada – UN [IAEA, Vienna]**: Attendance by Bob Truong of the Canadian Nuclear Safeguards Program at the 2004 Toronto Workshop led to collaboration with Canadian researchers. As a direct result, IAEA have adopted VITA [shown at that Workshop and other IST-059 meetings] for text mining and knowledge discovery in their next-generation enterprise decision support system, rollout in 2007. They will use VITA with special aim at tracking global traffic in nuclear weapons materials, devices and technologies.
- **Canada – WHO**: Developments in the visualisation program VITA engendered by collaboration between Canada and the UN are now used by the Global Public Health Information Network [GPHIN] which the Public Health Agency of Canada runs for the World Health Organization. VITA will be introduced to WHO-HQ [Switzerland] and to the USA Centers for Disease Control [CDC].
- **USA – Norway**: Awareness of the hypernode algorithm developed by Norwegian researchers has led to planning for implementation of that algorithm into multiple USA programs.
- **USA – Canada**: Awareness of Canadian work on information extraction from free text to extract network information has led to discussions between DRDC and Wave Technologies, a USA company under contract to the USA Army, for placement of the tool into the new Urban Warfare Analysis Center for testing by operational analysts.
- **USA – Canada**: Canadian and USA researchers have developed and are seeking funding for a program of work to answer the question of whether specific types of networks react in predictable ways to various influence factors. If such can be catalogued, predictive, planning and reactive tools can be developed to benefit coalition forces world over.
- **Norway – Denmark – Canada – USA**: Professors at Aalborg University requested RTG members' participation in the European Center for Counterterrorism Research and Studies.
- **GBR – Canada**: Researchers at QinetiQ Corp. [UK] and Health Canada initiated collaboration on mathematical tools for infectious disease outbreak management as a direct result of sessions and contacts at the RTG and N/X Workshops. They are working together developing support tools for the Canadian government pandemic response.
- **GBR – Canada**: QinetiQ and Health Canada are working together for dynamic network experimentations.
- **GBR – Norway**: QinetiQ and Norway are working together on hypernode techniques and network uncertainties.

8.2 VALUE ADDED

The value of the work of IST-059 can be seen both in the avoided costs of development that would have been incurred by the nations and in the value added in opportunity costs for present and future operations.

The IST and N/X developed several algorithms of general utility; the avoided costs to develop the algorithms commercially have been estimated at three million dollars American.

Beyond this, the meetings and collaborations have offered considerable added value, both at present and in the foreseeable future:

- Algorithms developed as a result of the group's interactions are already used for intelligence in nuclear counter proliferation and disease surveillance.
- The Framework will assist the development of visualisation systems for network-enabled operations, by introducing real-world considerations into network abstractions.
- Using the Framework to analyze a network within its embedding environment will assist in discovery of potentially important unobserved links, such as in and among terrorist groups. Embedding environments such as geographic location, social historical similarities, and other cross-domain embeddings can highlight areas of special interest.
- The Survey has elucidated many aspects of network display that prove useful in generalizing display technologies, using the Framework to describe user tasks in a common language.
- The RTG papers and reports have led to the creation of a potential DARPA program in the USA to treat dynamic and layered multimodal networks.
- The RTG has embarked on a program to develop a unified theory of networks which will encompass networks of all kinds, including social networks, physical networks, logistic networks, semantic or conceptual networks, and virtual networks; the unified theory will ultimately treat layered networks, network dynamics, information embedded generated within networks.

The following table shows the types of value added gained by each country by participation in this group:

Table 8-1: Value Added to Member Nations via IST-059 Group Activity

Gaining Country	Item	Effort Saved / Items Produced	Value
USA, GBR, DNK, CAN	Hypernode Algorithm	2 FTE for one year per country	$500,000 USD x 4 countries = $2,000,000 USD
USA, NOR, DNK, GBR	Visualisation Survey	80 hours of effort per country	$6,500 USD x 4 countries = $26,000 USD
USA, GBR, DNK, NOR, CAN	Knowledge of ongoing programs and lessons learned	960 hours of effort (.5 FTE for one year) per country	$125,000 USD x 5 countries = $1,125,000 USD
		TOTAL VALUE ADDED	**$3,151,000 USD**

8.3 BEGINNINGS OF THE UNIFIED NETWORK THEORY

Perhaps one of the most significant impacts of the work of IST-059 is the birth of a Unified Network Theory which was spawned from the need to connect a user’s network problem in laymen’s terms with the analytical and display tools in various software packages. In order to connect the two and make a good match, it was necessary to prepare a type of middle ground language that could act as a translation mechanism. The IST-059 Framework, associated taxonomy and information theoretic analyses resulted. Although this translation mechanism will be highly successful in allowing us to map user needs to available software packages; the truly amazing impact comes when you realize that this mechanism is in fact the beginnings of a Unified Network Theory. Specifically, that it lays the foundation for a complete theory into which any network scenario in any domain may be mapped and within which we can operate on, analyze, and visualise any network in any domain without bounds. The potential application of the Unified Network Theory and the implications of these applications are profound. As a separate project, the Group recommends that this theory be completed and applied.

Chapter 9 – CONCLUSIONS AND RECOMMENDATIONS

The following summarizes the conclusions and recommendations from each of the Framework and Survey results, the workshops, and the impact made by the Group's efforts.

9.1 FRAMEWORK

9.1.1 Conclusions

Many different kinds of network representation have been developed, but without a coherent foundation that would allow good representations to be used for other projects. A good Framework provides that foundation.

- A good representation supports the purposes of a user effectively.
- A Framework requires consideration of both the user and the range of network properties that might be represented in support of the user's purposes. Therefore a Framework must consider the nature of real networks as well as the properties of abstract mathematical graphs.
- Real networks are more complicated than are the abstract mathematical networks, though the mathematics remains relevant to the real networks.
- Real networks are often fuzzy. Links and nodes may be of variable quality. Nodes transform the kinds of traffic they receive and emit.
- Real networks are embedded in user-relevant context that affects their properties and behaviour. The context may itself be a network.

Within the ambit of IST-059/RTG-025, the following steps seem necessary, though IST-059/RTG-025 does not have the resources to complete them:

- Complete the Framework by:
  - Categorizing computable network attributes.
  - Categorizing Network-related user tasks.
  - Categorizing network-related display techniques.
  - Develop mappings across categorizations:
    - Task – attribute; and
    - Attribute – display.
- Incorporate interaction (*the theme of the follow-on RTG*).
- Link the Framework with the Survey of Network Visualisation Software.
- Describe the Framework process for end-users.
- Propose support software to guide the user in the Framework process.
- Test Framework use in different scenarios, and rework.
- Publish for general use.

9.1.2 Recommendations

A follow on Technical Team to be known as IST-085 has been recommended to the RTB by the IST Panel for a start in January 2009. The follow-on group will have the title “Interactive Visualisation of Network Dynamics”. Many studies have shown that allowing the user to have hands-on control of a display usually enhances the user’s understanding of the display. Activities of this group will include:

- Compare the utility of various interactive visualisation styles for providing the user knowledge of the dynamics of a network and subsequent trends.
- Develop the required experiments to provide insight into what characteristics of interactive visualisations are most likely to aid the military user in determining and predicting the types of change happening within a network, given various influence factors.
- Produce a report highlighting interactive visualisation methods that facilitate and make more effective the analysis of network dynamics in applications such as netcentric warfare, counterterrorism including bioterrorism, peacekeeping, public security, and peace support operations.
- Background study will include collecting and analysing information about the state of the art for such visualisation in various nations across various problem domains, and integrating/synthesizing the state of extant technology to:
  a) Formulate the experimental designs,
  b) Extend the network dynamics capability of the Framework and Survey of IST-059/RTG-025 to aid in generating new concepts for displaying and interpreting network dynamics, and
  c) Develop recommendations for use in future network visualisation systems.
- If feasible, the Group will mount a demonstration giving the opportunity for hands on exploration of interactive visualisations of network dynamics to show the experimental design and reported results.

9.2 SURVEY

9.2.1 Conclusions

In the survey, four areas of focus were identified as being required to advance the network visualisation field.

- **Information Sharing Support**, which includes theory, standards, and software, is needed to allow researchers from diverse application domains to work together. Working together across disciplines will enhance creativity.
- **Network Representations** must be improved to provide satisfactory presentations of large and/or dynamic networks, along with an indication of uncertainty and adaptable to specialized hardware.
- **Decision Support** is the end goal of displaying the data to the user, and the special properties of networks must be exploited to assist the user in accomplishing their task. Prediction of future network behaviour is an unaddressed research area.
- **Evaluation** must be integrated into the research process. If a method is to be accepted by the community, good science requires proof that the method satisfies a human user. If a method is to be transitioned into a commercial product, industry requires proof of its efficacy.

9.2.2 Recommendations

Network visualisation is a fairly new discipline and its foundation is still to be defined and accepted by the scientific community. Advances in the domain of information visualisation in term of standards, representations, and evaluations will necessarily benefit network visualisation. Building on the good work already done and standardizing the evaluation process will better focus our efforts.

9.3 WORKSHOPS

9.3.1 Conclusions

Most network information sets do not take “time” into account.

Network analysis and visualisation are:

1) Important tools in the fight against terrorism; and
2) Useful for tracking disease and attacks on computer systems (virus, etc.).

Better view of the networks and what is occurring on them is needed. Multiple views of the network are needed:

- A logical view that shows communications links, routers, servers, firewalls and applications.
- A physical view that overlays the logical view on a geographical representation.
- A transactional view that shows if the various applications are functioning. Is logistics delivering just in time supplies? Are invoices being paid on time?
- An operational view that shows commanders and staff are able to use the networks to gain the advantage that Network Enabled Ops promises. Network staff can prioritize restoral on the basis of operational priorities.

There is not enough collaboration between the academic researchers and the defence community. The academic researchers would like realistic information to test their systems. If the systems are not tested on realistic information they might not be developed into useful systems for defence.

There is a critical need for advancing network Command and Control.

Progress has been marginal in the Network Operations Centres.

Two of the topics chosen by the workgroups “Reliability and uncertainty in situation awareness of Network Visualisation” and “Vulnerability and Risk Assessment” are important topics needing more research.

9.3.2 Recommendations

Contact between users, developers and researchers should be encouraged.

The problems need to be defined to be solved in ways that even civilian researchers may work on them. Better ways to get laundered data to researchers/modellers should be established.

Operational studies and analyses of visualisation needs from the analyst’s and commander’s viewpoints should be initiated. A seamless environment across Net Ops Centres and R&D Labs should be created and R&D results should be used in the Ops Centres as soon as possible.

One "low hanging fruit" is to use higher resolution display technology or an appropriately matrixed array of high resolution displays oriented in such as manner as to maximize the information availability without "overloading" the user. Further research may improve how displays of this calibre can be organized to maximize the information output without creating an overload situation.

Experimentation with 3-D visualisation should be encouraged.

More research on "Reliability and uncertainty in situation awareness of Network Visualisation" and "Vulnerability and Risk Assessment" should be encouraged.

Concerning uncertainty and reliability the following is necessary:

- A clear definition of reliability and uncertainty is needed.
- Development of visualisation concepts and prototypes, defining what uncertainty and reliability conveys.
- Conduct experiments with representations of uncertainty and reliability.
- Development of consistent techniques for determining uncertainty and reliability.
- Development of intuitive techniques for visualising uncertainty and reliability.

9.4 SIGNIFICANCE OF THE WORK

The work of IST-059 has had significant **impact and influence on international collaboration, future research and ongoing programs**. Perhaps one of the most significant impacts of the work of IST-059 is the birth of a **Unified Network Theory** which was spawned from the need to connect a user's network problem in laymen's terms with the analytical and display tools in various software packages. In order to connect the two and make a good match, we were forced to prepare a type of middle ground language that could act as a translation mechanism. The framework, associated taxonomy and information theory resulted. Although this translation mechanism will be highly successful in allowing us to map user needs to available software packages; the truly amazing impact comes when you realize that this mechanism is in fact the beginnings of a Unified Network Theory. Specifically, that it lays the foundation for a complete theory into which any network scenario in any domain may be mapped and within which we can operate on, analyze, and visualise any network in any domain without bounds. The potential application of the Unified Network Theory and the implications of these applications are profound. As a separate project, the Group recommends that this theory be completed and applied.

Annex A – FORMAL DOCUMENTS: TAP, ToR, PoW

Technical Activity Proposal (TAP)

<table>
<tr><td><i>ACTIVITY</i></td><td>RTG</td><td colspan="11" rowspan="2">Visualisation Technology for Network Analysis</td><td colspan="3">04/2004</td></tr>
<tr><td><i>REF. Number</i></td><td>IST-059/RTG-025</td><td colspan="3">01/2005</td></tr>
<tr><td colspan="2"><i>PRINCIPAL MILITARY REQUIREMENTS</i></td><td>2</td><td>3</td><td>4</td><td></td><td></td><td></td><td></td><td></td><td colspan="3">NS</td><td colspan="3">12/2007</td></tr>
<tr><td colspan="2"><i>MILITARY FUNCTIONS</i></td><td>1</td><td>2</td><td>3</td><td>6</td><td>7</td><td>9</td><td>11</td><td>12</td><td>13</td><td>14</td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="2"><i>PANEL AND COORDINATION</i></td><td colspan="8">IST: Information Systems Technology</td><td colspan="3">--</td><td colspan="3">--</td></tr>
<tr><td colspan="2"><i>LOCATION AND DATES</i></td><td colspan="11">Participating Nations: semi-annually until</td><td colspan="3">P-I</td></tr>
<tr><td colspan="2"><i>PUBLICATION DATA</i></td><td colspan="7">TR</td><td colspan="3">2007</td><td colspan="2"></td><td colspan="2"></td></tr>
<tr><td><i>KEYWORDS</i></td><td>Network analysis</td><td colspan="5">Social networks</td><td colspan="5">Computer networks</td><td colspan="4">Data overload</td></tr>
<tr><td colspan="2">Network centric warfare</td><td colspan="5">Situation assessment</td><td colspan="5">Human-machine</td><td colspan="4">Massive datasets</td></tr>
<tr><td colspan="2">Multimedia visualisation</td><td colspan="5">Command and Control</td><td colspan="5">Link analysis</td><td colspan="4">Counterterrorism</td></tr>
<tr><td colspan="2">Knowledge Discovery</td><td colspan="5">Information</td><td colspan="5"></td><td colspan="4"></td></tr>
</table>

I. BACKGROUND AND JUSTIFICATION

During the course of the work of the IST RTG-002 and RTG-007 Technical Teams, reinforced by the deliberations and recommendations of the Quebec and Halden Workshops, and supported by the observations of the 2001 and 2003 meetings of the visualisation Network of Experts (N/X), network representation and analysis issues appeared repeatedly in many guises across numerous problem domains. There is a need to understand what visualisation technologies to use and how to use them effectively to support network discovery and analysis tasks. In this context, networks include both "structural" networks – e.g. information and service infrastructural networks – as well as "logical" networks – e.g. social networks which show the organizational relationships among their elements. Such social networks might show, for example, the relationships among terrorist cells and their members or the historical relationships among international or local agreements and laws.

Anticipated military benefits include a better understanding of available visualisation technology and techniques and their potential uses and benefits as applied to military and intelligence network analysis tasks. Visualisation methods to facilitate and speed the analysis of networks in uses such as netcentric warfare and counterterrorism, and peacekeeping and peace support operations would be considered. The exchange of scientific and technical information among member nations will be ongoing throughout the life of the RTG.

II. OBJECTIVE(S)

Produce a document to further understanding of visualisation technology and techniques as applied to network analysis tasks in order to help identify where and how visualisation methodology can realistically benefit such tasks. This will involve collecting and analysing information about the state of the art in network data visualisation in various nations across various problem domains, and integrating/synthesizing the state of extant technology to a) generate new concepts for displaying and interpreting network data, b) develop recommendations for use in future network visualisation systems, and c) identify future research issues that must be addressed to advance the field. This will include using visualisation technology to discover relationships, present relationships and to analyse relationships within and across both structural and social networks. Comparative evaluation of the effectiveness of display concepts/techniques within a given role will also be considered.

III. TOPICS TO BE COVERED

- Visualising Networks and Network Data.
- Extraction/Discovery, analysis, representation and evaluation of social network data from reports, messages and other documents.
- Applications to Situation Awareness and Decision Support:
  - Applications to Network-Centric Warfare.
  - Optimising Human-Machine Interface for Network Data.
  - Evaluation Methods and Tools.
  - Technology Overview and Review (Includes forecasting what military needs will and will not be met in Private and University sectors).

IV. DELIVERABLE

1) Workshop on "Visualising Network Data" (2006) Note: The actual workshop title and workshop scope will be developed by the RTG during its first year of operation.

2) Final Report (2007).

V. TECHNICAL TEAM LEADER AND LEAD NATION

The Technical Team Leader is Vincent Taylor (CAN). The Lead Nation is CAN.

VI. NATIONS WILLING TO PARTICIPATE

CAN, DEU, DNK, GBR, NOR, ROU, USA.

VII. NATIONAL AND/OR NATO RESOURCES NEEDED

Subject matter experts and technical specialists from the participating nations and organizations. Participants will need access to the Internet, since it is intended to interact electronically among the experts.

VIII. RTA RESOURCES NEEDED

Support for two Consultants per year.

Terms of Reference (ToR)

RESEARCH TASK GROUP ON

“VISUALISATION TECHNOLOGY FOR NETWORK ANALYSIS”

IST-059 / RTG-025

I. ORIGIN

A) Background

Visualisation, a means by which people make sense of complex data, can be seen as a human activity supported by technology. A key element of visualisation is the interface through which the human interacts with the data. It includes both the “how” as well as the “what, when, where, and why” of information presentation and control. Visualisation technologies include search engines, algorithmic processes, display and control devices, but what matters is how these technologies enhance and allow people to do their tasks in a timely and effective manner.

During the course of the work of the visualisation Technical Teams, IST RTG-002 and RTG-007, network representation and analysis appeared repeatedly as issues in many guises across numerous problem domains. This was reinforced by the deliberations and recommendations of the Quebec and Halden Workshops and was also supported by the observations of the 2001 and 2003 meetings of the visualisation Network of Experts (N/X).

There is a need to understand what visualisation technologies to use and how to use them effectively to support network discovery and analysis tasks. In this context, networks include both “structural” and “logical” networks. A structural network would include the classic networks such as computer networks, railway networks, gas and oil distribution networks, etc., all of which have a basic structure, have the concept of routing or switching, and support applications – e.g. e-mail over a network of fibre optic, satellite and radio links; the dispatch and movement of trains over railways; the movement of gas and oil through pipelines, etc. A logical network – e.g. a social network which maps the relationships among its elements – might reveal, for example, the organization among terrorist cells and their members, the historical relationships among international or local agreements and laws, or the propagated effects of accidental or deliberate damage to elements of inter-related infrastructures.

Although there appears to be a large difference between the two types of network, from a visualisation perspective they have much in common. Each has layers of interest to the observer (visualiser) which include a) its structure, whether physical as in the case of a network of computers and communication elements, or logically inferred, as in the case of a network showing the relationships in a hierarchy among the members of a community of interest; b) its potential behaviour – i.e.. the activities that may take place over the network and which can be measured or inferred by their influence on the network; and c) its current and predicted behaviour, along with its effect on items of interest to the observer.

It is the belief that, although the two types of network are normally within the domains of different user communities – e.g. managers vs. analysts, the visualisation technologies that need to be employed have much in common and cross-fertilization among the communities would provide mutual benefit over the longer term.

B) Justification (Relevance for NATO)

Anticipated military benefits include a better understanding of available visualisation technologies and techniques with respect to their potential uses and benefits in military and intelligence network analysis tasks. Visualisation methodology to facilitate and speed the analysis of networks in uses such as network centric operations, counterterrorism, peacekeeping and peace support operations will be considered.

Networks of relationships include causal or probabilistic networks that affect planning of military operations. These relationships are often not made evident in current planning systems but are created in a commander's mind through his experience and interpretation of his map displays and related data presentations. Improvements in the display of such relationships should promote common understanding across roles as well as improving the speed and robustness of operational planning.

II. OBJECTIVE

The area of research is the enhancement of human ability to visualise the networks with which they are concerned. This will include the understanding of the application of visualisation technology and techniques to network management and analysis tasks in order to help identify where and how such methodology can realistically benefit the human performing such tasks. To do this will involve collecting and analysing information about the state of the art in network data visualisation across various problem domains, perform experiments to integrate/synthesize promising current technology to determine its capabilities for supporting network analysis, and identifying areas in which further research would be profitable. The exchange of scientific and technical information among member nations will be an ongoing background activity throughout the life of the RTG. Suggested application domains include situation awareness and decision support for netcentric and counter-terror operations.

1) The RTG will produce a report that will identify and categorize visualisation technology and techniques that can be applied to network analysis tasks. This document will help military users to identify where and how visualisation methodology can realistically benefit their tasks.

   Research topics include:

   - Representation of network structure and activity;
   - Extraction/Discovery, analysis, representation and evaluation of network data from reports, messages and other documents;
   - Human-Machine Interface for network data;
   - Evaluation methods and tools; and
   - Technology overview and review (Includes forecasting what military needs will and will not be met in Private and University sectors).

   The report will document members' experiments carried out to support the goals of the RTG. Such experiments may involve one or more nations.

2) In support of its objectives, the RTG expects to develop and deliver a workshop during its second year on "Visualising Network Information", or similar topic, at a location to be determined.

3) The RTG will continue to foster the activities of the visualisation Network of Experts originally created by RTG-002.

4) The RTG will have a three year term.

III. RESOURCES

A) Membership

The membership will be research and military experts from the nations and NATO agencies who have experience in network and visualisation technologies and/or relevant defence applications.

The Lead Nation will be CAN.

The Technical Team Leader will be Mr. Vincent Taylor.

Nations that have agreed to participate in the RTG are: CAN, DEU, DNK, GBR, NOR, ROU, USA.

B) National and/or NATO Resources Needed

Subject matter experts and technical specialists from the participating nations and organizations are required. Participants will need frequent access to the Internet, since it is intended to interact electronically among the experts. The RTG is expected to meet twice each calendar year in the member nations. Nations should be prepared to fund the travel for their delegate(s) to participate in all meetings. Each nation should be prepared to host at least one meeting during the lifetime of the RTG.

Participants should be prepared to loan technology to RTG members to allow for approved collaborative or cooperative experimentation supporting the Program of Work. Such technology would be controlled and remain the property of the donor country.

C) RTA Resources Needed

Two consultants per year.

IV. SECURITY CLASSIFICATION LEVEL

The RTG may operate up to NATO SECRET.

V. PARTICIPATION BY PARTNER NATIONS

The RTG will be open to Partner Nations.

VI. LIAISON

NC3A, NATO Transformation Command – liaison to understand the operational research needs.

HFM Panel – liaison to maintain a human factors awareness with respect to cognitive issues.

IST C2 Technical Teams (TT) – liaison to ensure minimal overlap in the programme of work, as well as to coordinate the promulgation of relevant technological results that might impact the work of the TTs.

RTO ToR FORM – NOVEMBER 2001

Programme of Work (PoW)

RESEARCH TASK GROUP ON

"VISUALISATION TECHNOLOGY FOR NETWORK ANALYSIS"

IST-059 / RTG-025

BACKGROUND

Visualisation, a means by which people make sense of complex data, can be seen as a human activity supported by technology. A key element of visualisation is the interface through which the human interacts with the data. It includes both the "how" as well as the "what, when, where, and why" of information presentation and control. Visualisation technologies include search engines, algorithmic processes, display and control devices, but what matters is how these technologies enhance and allow people to do their tasks in a timely and effective manner.

During the course of the work of the visualisation Technical Teams, IST RTG-002 and RTG-007, network representation and analysis appeared repeatedly as issues in many guises across numerous problem domains. This was reinforced by the deliberations and recommendations of the Quebec and Halden Workshops and was also supported by the observations of the 2001 and 2003 meetings of the visualisation Network of Experts (N/X).

There is a need to understand what visualisation technologies to use and how to use them effectively to support network discovery and analysis tasks. In this context, networks include both "structural" and "logical" networks. A structural network would include the classic networks such as computer networks, railway networks, gas and oil distribution networks, etc., all of which have a basic structure, have the concept of routing or switching, and support applications – e.g. e-mail over a network of fibre optic, satellite and radio links; the dispatch and movement of trains over railways; the movement of gas and oil through pipelines, etc. A logical network – e.g. a social network which maps the relationships among its elements – might reveal, for example, the organization among terrorist cells and their members, the historical relationships among international or local agreements and laws, or the propagated effects of accidental or deliberate damage to elements of inter-related infrastructures.

Although there appears to be a large difference between the two types of network, from a visualisation perspective they have much in common. Each has layers of interest to the observer (visualiser) which include a) its structure, whether physical as in the case of a network of computers and communication elements, or logically inferred, as in the case of a network showing the relationships in a hierarchy among the members of a community of interest; b) its potential behaviour – i.e.. the activities that may take place over the network and which can be measured or inferred by their influence on the network; and c) its current and predicted behaviour, along with its effect on items of interest to the observer.

It is the belief that, although the two types of network are normally within the domains of different user communities – e.g. managers vs. analysts, the visualisation technologies that need to be employed have much in common and cross-fertilization among the communities would provide mutual benefit over the longer term.

MAJOR WORK ITEMS

A) General Overview

The area of research is the enhancement of human ability to visualise the networks with which they are concerned. This will include the understanding of the application of visualisation technology and techniques to network management and analysis tasks in order to help identify where and how such methodology can realistically benefit the human performing such tasks. To do this will involve collecting and analysing information about the state of the art in network data visualisation across various problem domains, perform experiments to integrate/synthesize promising current technology to determine its capabilities for supporting network analysis, and identifying areas in which further research would be profitable. The exchange of scientific and technical information among member nations will be an ongoing background activity throughout the life of the RTG.

Recommended principal application domains to be considered include situation awareness and decision support for netcentric operations and for counter-terror operations.

B) Work Items

1) Plan overall activities of the RTG:

   This task will: confirm expected national or agency contributions in terms of manpower, potential data, models, testbeds, targets, equipment, computer time, etc., available to support the group; validate the work items for the RTG, given the resources expected to be available from the nations; refine the PoW and provide the appropriate bounds and milestones on the work items; and will confirm the modus operandi of the group, including agreement on hardware and software to be used for editing reports from the group.

2) Survey visualisation technology of potential relevance in network analysis:

   This will include technology that is in production use as well as technology that is in the research and development stage within the individual countries. The RTG will attempt to identify in a gross manner how the technology addresses the visualisation of one or more of: network structure; potential network behaviour; and actual and predicted network behaviour, particularly in the agreed application domains.

3) Identify and categorize promising technologies:

   This task will require the RTG to agree on what technologies appear to have the most merit to address network analysis problems. Experimentation, either by individual countries or in collaboration by two or more partners may follow in order to validate findings and to discover or confirm the essential characteristics of the technology. The nature of any experimentation will depend on the technologies chosen.

4) Develop and produce a Workshop:

   In support of its objectives, the RTG expects to develop and deliver a workshop in October 2006 on "Visualising Network Information" (IST-063/RWS-010) in Copenhagen.

5) Define a network visualisation framework:

   This task is to initiate the development of descriptive and functional frameworks for network visualisation. A descriptive network visualisation framework will enhance an understanding of the commonalities of different ways of presenting network properties so that methods appropriate to one can be transferred to another. A functional network visualisation framework will characterize the interaction of a human operator with the network representation.

6) Support the visualisation Network of Experts (Vis N/X):

   The RTG will continue to foster the activities of the visualisation Network of Experts originally created by RTG-002. The RTG believes that the Network of Experts will produce workshops in 2005 and 2007 that support the program of the RTG.

7) Produce Final Report:

   The RTG will produce a final report that will summarize the activities during the life of the RTG. The report will document experiments carried out to support the goals of the RTG. It will identify and categorize identified visualisation technology and techniques that can be applied to network analysis tasks, particularly within the agreed application domains. It will address at least the following topics:

   - Technology overview;
   - Representation of network structure and activity;
   - Uncertainty, validity and reliability;
   - Scalability issues;
   - Extraction/Discovery, analysis, representation and evaluation of network data from reports, messages and other documents;
   - Human-Machine Interface for network data;
   - Evaluation methods and tools; and
   - Technology forecast, particularly with respect to what military needs will and will not be met in Private and University sectors).

   This document will help operational users to identify what visualisation methodology could realistically benefit their tasks and where and how it should sensibly be used.

SIGNIFICANT MILESTONES

October 2006 – Workshop delivery (IST-063/RWS-010).

December 2007 – Final report.

NATIONS PARTICIPATING

The membership will be research and military experts from the nations and NATO agencies who have experience in network and visualisation technologies and/or relevant defence applications.

The Lead Nation will be CAN.

The Technical Team Leader will be Mr. Vincent Taylor.

Nations that have agreed to participate in the RTG are: CAN, DEU, DNK, GBR, NOR, ROU, USA.

CONFIRMED NATIONAL OR AGENCY CONTRIBUTIONS

TBD during RTG Planning Phase.

CONTACT INFORMATION

Team leader:

V.K. Taylor
DRDC Ottawa
3701 Carling Avenue
Ottawa, Ontario K1A 0Z4
CANADA
Vincent.Taylor@drdc-rddc.gc.ca
+1 (613) 993-9946 voice
+1 (613) 993-9940 fax

Annex B – THE IST-059 FRAMEWORK FOR NETWORK VISUALISATION

M.M. Taylor and A.K.C.S. Vanderbilt

B.1 INTRODUCTION

Why might a user want to visualise a network, and what about the network might she want to visualise? The concept of a network pervades so many different areas, including the networks of influence in the genetic processing in a cell, the detection of intrusions in computer networks, the analysis of vulnerabilities in the electricity supply system, the discovery of key personnel in terrorist groups, the interplay of ideas in scientific discovery, the dynamics of planning a military operation, the optimization of routings of supplies for disaster relief, the effects of changes in traffic light timings or in street signage, the effects of gossip and of advertising on the growth of disease epidemics, and so on and so forth. All these, and many more equally varied applications, intrinsically embody networks and their visualisation. A Framework for network visualisation should encompass all of these possibilities without being so diffuse as to be useless for any particular problem.

The IST-059 Framework for Network Visualisation (Chapter 2) is intended to help a user who is concerned with some problem involving a network to clarify the problem and to find appropriate tools to solve the problem. It is a process as well as a set of categorizations of networks, of tasks, of perceptual modes, of display techniques, and of data types. The categorizations are linked, and together they help specify what kinds of tools might be useful.

The Framework is useful on its own, but should be more useful when implemented in software as a front-end to the IST-059 Survey database of available network applications and tools (Chapter 3). The integration of the Framework with the Survey is addressed in Chapter 5. IST-059 did not address software implementation of the framework, nor its interface with the Survey, considering those matters to be more than could be addressed in the lifetime of the group. In this Annex, the Framework is considered as a stand-alone construct.

B.1.1 Framework Concept

One concept important for real-world networks that does not occur with graphs is the “embedding field” of a network. We have not encountered this concept elsewhere. The notion of “Framework” means different things to different people. To IST-059, a Framework is like the skeleton of a body; it provides the linkages among disparate components and makes possible a process for getting things done. More specifically, the Framework uses a controlled series of questions to help the user clarify the issues that arise in the task at hand, using taxonomies of the different components that should be considered. This clarification helps the user to decide how best to display the available data, and, in conjunction with a survey of network analysis and display tools, to decide what software might serve the task at hand.

It is important to note that IST-059 is concerned not just with abstract networks in the form of mathematically tractable graphs, but with networks that appear in real world problems. In such problems, the context of the network may be as important for the user’s understanding as the network itself. The representation of real-world networks in context is a rather richer domain than that of graphs, although graph theory is nevertheless applicable to many issues of real networks. The mathematical analysis of networks is discussed in Annex C.

B.1.2 Framework and Survey

In parallel with the development of the Framework, IST-059 also conducted a Survey of applications and software tools that aid the analysis and visualisation of networks. Although the Survey stands on its own, as can the Framework, it was realized that the Framework could be implemented in the form of an interface, or front-end, to the Survey database. We do not consider the integration of the Framework with the Survey in this Annex, since no implementation was attempted by IST-059, and because the conceptual basis of integration has a chapter of its own in this Report (Chapter 5).

B.1.3 Why Create a New Framework?

Representations of networks come in a great variety of forms, each designed to show off some aspect of the network, and each based on the thoughts and intuitions of some designer. Nearly six hundred very different examples produced for a wide range of applications were illustrated at the "Visual Complexity" Web-site http://www.visualcomplexity.com/vc/ as of the beginning of June 2008.

Networks have many different properties that the user might find important for the purpose of the moment, and many different approaches have been taken to creating representations that support the user's ability to visualise these properties. Clearly, the designer of any one representation will have been influenced by previous ideas, but it is not easy for someone who wants to design a representation for a particular application to generalize from earlier examples to the new case, unless the kind of network and the needs of the user are clearly analogous in the two situations.

The reason for creating any representation is to aid the human to visualise some aspect of the thing represented. The computer presents the data in some form, perhaps pictorial, perhaps not. Based on that presentation and his or her background memory and skill, the human user visualises its implications as one of the routes to understanding the data. The computer's presentation of the data is an aid to the user's visualisation, not its content.

Visualisation is one way people try to understand situations. Logical analysis acts in concert with visualisation as a parallel route to understanding. The kinds of display that support logical analysis may well differ from those that support effective visualisation, in that analysis is helped by individuating items and making their connections explicit and distinct, whereas visualisation is often aided by more diffuse global representations. Although the two normally work together, this report concerns only visualisation. Displays that aid visualisation may often be used in conjunction with displays that support logical analysis.

The content of the user's visualisation incorporates not only the data displayed by the computer, but also material from the user's memory and imagination. This important point is often lost in the design of displays, many of which are designed to show as much of the data as is reasonably possible, including things the user might be expected to know already. Experts and novices often see different things in a display, especially a complex one.

The user has some purpose in wanting the data to be displayed. Perhaps the purpose is no more than idle curiosity, but more commonly it is in support of some task of the moment. At any particular moment, the user's purpose probably does not require very much of the available data to be presented, but whatever data are presented must have context, whether it be supplied in the presentation or by the user's memory and imagination.

The objective of a Reference Framework is to provide a guide to assist the user to find the most suitable display for the task at hand, and to aid generalization from one situation to another. It should also assist in both the design process for new kinds of display and the evaluation of a completed design. A good Reference Framework, in conjunction with a Survey of extant applications and software tools, should also assist a user to determine whether some particular software is likely to be useful for a particular purpose, and to guide the selection of the most useful out of a collection of visualisation system.

In summary, the reasons for trying to develop a Framework for Network Representation are:

- Numerous ad-hoc examples of network representations have been created for specific applications, some of them very good for their purpose. The Framework should help the user determine which, if any, of these tools should be expected to be useful for the current task.
- It is usually not clear how the insights that led to particularly effective representations of some particular network can be generalized to new situations. A good Framework should help isolate the conditions for which different insights are helpful.
- Users need to see different aspects of network structure and function, and many of those aspects are not well served by extant representation techniques; a Framework may help inspire new modes of representation.

B.1.4 Intellectual Background of the IST-059 Framework

The Framework has several disparate roots, some of which come from the work of the predecessor groups of IST-059 (DRG Panel 3/RSG-30, IST-013, and IST-021, collectively known, along with IST-059, as “VisTG”). The VisTG Reference Model for visualisation (Annex H) is one of these roots. The taxonomies of data types and display types presented in the Final Report of IST-013 [1] provide another. A third separate starting point is the RM-Vis framework developed by TTCP C3I AG2 (described in Annex G). These disparate intellectual starting points merge and are extended in the Framework development.

B.1.4.1 Thinking about Representation: Abstraction

At the IST-043 Workshop in 2005 [3], Working Group 5 produced the following definition of a network: *A Network is an array of nodes that exchange “stuff” over links on containers under a certain protocol and following a determined path*. The WG-5 definition does not distinguish between functioning networks in the real world and their abstract representations, although it seems to lean more toward the real world. Nevertheless the definition works well for many networks.

In this Annex, “stuff” is called “traffic”, which may flow continuously or in discrete packets between nodes. In some networks traffic is conserved, in which case a transmitting node loses what passes along a link to a receiving node, which then gains the transmitted traffic. This would be the case, for example, of cars travelling between places in a road network. In other networks, traffic is not conserved, and a transmitting node does not necessarily lose what a receiving node gains. For example, a person transmitting some item of knowledge to another does not lose what the other gains. Both of these examples are of traffic-bearing networks, but many networks of interest bear no traffic, in contradiction to the WG-5 definition. The network of friendship relationships among a group of people is one that exchanges no “stuff” over its links (though the people concerned may do so).

Abstract representations of networks are the subject of mathematical graph theory and Social Network Analysis (SNA, see Annex C). Mathematical graph theory is valuable for network analysis in the same way that stress analysis of steelwork is valuable in the construction of bridges and buildings. Sometimes one can get away without it, but usually the graph-theoretic work is important to the final result. All the same, just as the stress analysis of a bridge or building tells little about how the structure will work on its site or how it will be viewed aesthetically, so the graph analysis can seldom be sufficient to provide a feeling of how the real world network works in its natural context.

Considering only the network itself and ignoring its context, there is a range of possible levels of abstraction. At one end of the range, nodes have only the topological property of being vertices at which connecting links meet. At the other end are the full-blown real-world conditions in which traffic emerging from a node on one or more links may be qualitatively distinct and temporally separated from traffic entering that node over other links, and may be influenced by the effects of the context (e.g. stray capacitances for the network of connections in an integrated circuit, distracting sights or events in a social network). Real-world nodes should be treated as processors, and sometimes so should links.

Intermediate levels of abstraction are possible. The mathematical representation might, for example, treat a node as emitting traffic only when a number of its input links have been active (e.g. a Petrie Net), or a link might be represented as being capable of holding a limited quantity of traffic at any one moment, or as imposing a probabilistic delay upon traffic emitted by its tail node before the traffic is delivered to its head node (as might be the case for a representation of real traffic flow over a road network). Networks in the conventional analysis of System Dynamics are at this level of abstraction. Nodes and links might be of various characters such that nodes of type A communicate only with nodes of type B. There are many possibilities at intermediate levels of abstraction.

Except in the most abstract case of simple topological representation, a network as a whole should ordinarily be regarded as a processing system. Whether its processing aspects and the associated dynamics are important to the user will determine how the network should be presented or displayed. If, for example, the real network is composed of computers and their interconnections, the processing at the nodes is potentially of unlimited complexity, but the network representation may abstract only the properties of specific chunks of data relevant to particular intercommunication protocols, ignoring all the other processing that might condition the use of those protocols. So there are levels of abstraction not only in representing the global properties of real networks, but also in representing the functional properties of the elements.

Most of the commonly considered properties of networks, such as centrality or cyclicity, refer to the topology of the network. In the real world, interest is often centred on the dynamical properties of the network, which may be constrained by the topology (e.g. oscillation cannot occur unless the network contains cycles), but which cannot be analyzed using only the topological level of abstraction.

Another dimension in which network representation is often abstracted is the fuzzy-crisp dimension. In the real world, links vary in quality rather than either existing or being absent, but in most graph-theoretic representations, two nodes either are or are not connected by a link, which may have a weight parameter, but for which the existence is all-or-none. In the real world, what the user wants to understand from the link may well determine the quality of the connection. Two users may see the same physical structure quite differently, and this difference can potentially be captured if the inherent fuzziness of the network is not abstracted away in an attempt to achieve topological purity.

B.1.4.2 Thinking about Representation: Structure

When a user needs to visualise something about a network, it seldom involves the entire network. Rather, the user may want to determine something specific about it, such as “Who is probably the leader of that group?”, “Where is the most vulnerable node?”, “How many different routes are suitable for this kind of traffic?”, “Where are the dangerous places?", "Is the threat increasing of diminishing?", and the like. Questions of this type are, in principle, answerable by a very low bandwidth information channel. For the example questions, answers might be, respectively "Joe Smith", "AZ175", "3", "Here, here, and here (pointing to a map)", "slightly increasing". Even if the network itself is very complex, it takes very few bits of information to convey the wanted information, and it is those few bits that the user must extract from what may be a complex display of an even more complex real-world situation.

The few bits that the user wants to extract by visualising a complex situation are elements of the structure of the situation. "Structure" was defined in information-theoretic terms by Garner [1], as the difference between two measures of uncertainty or entropy. The terms "uncertainty" and "entropy" are formally identical, though "uncertainty" implies that the quantity refers to something in someone's mind, whereas "entropy" implies that the quantity is a property of the physical system being quantified. Both refer to a summation of terms of the form $-p \log_2 p$ where p is the probability that some element of the system would take on the value it has. When talking about "uncertainty" that probability is in the mind of the observer of the system; when talking about "entropy" it is obtained from some physical measure. For example, if a network node Q has five links to other nodes and there are 100 nodes in the whole network, the probability that node Q is connected to an arbitrary other node is 0.05 if no other information about the network is available, and the observation that it is actually connected to node T provides $-\log_2 0.05 \approx 4.3$ bits of information.

The uncertainty about the connections of Q is 4.3 bits less after the connection to T is known than it was before that observation. Alternatively, the entropy of such a network in which node Q has a fixed connection to node T is 4.3 bits less than is the entropy of a similar network in which that connection is not fixed. That 4.3 bits represents the quantitative measure of the structure induced by fixing the Q-T connection. If one thinks of "uncertainty", the structure is structure known to an observer who previously knew that Q had five connections but not where those connections led. If one thinks of "entropy" it is structure in a network in which node Q could have exactly five connections. The mathematics is the same, but the implications are different.

When one is concerned with a user's ability to answer specific questions about a complex real-world situation, one must be concerned with the channels by which the necessary few bits of information are transmitted to the user, and in particular with the preservation through those channels of the structure that provides the answer to the user.

The entropy of the real world is very large, much too large to be accommodated in any computer dataspace, even when allowance is made for the structures inherent in the temporal and spatial correlations among the elements of the world. Sensors select what is entered into the dataspace without reference to the user's needs of the moment (unless the user interactively guides the sensors and controls the algorithms for selection). Accordingly, the entropy of the representation in the dataspace is also very large compared with the structure of interest to the user.

The display normally is of much lower entropy than is the entire dataspace, but if it is to be useful, it must contain the structure that the user wants to see in the real world. Numerically, this means that the same number of structural bits must pass through the channel real-world to dataspace to display, while the total bits implicit in the entropy of those environments is drastically reduced. In effect, the transition between stages acts as a filter, and that filter must be matched to the user's requirements. A radio filters from the airwaves one station, but the user interactively tunes to the wanted station. So likewise the user may control interactively the selection and algorithms that relate the display to the dataspace. "Tuning" implies the ability to guide the information channel to sustain the structure that will allow the user to answer the question.

The following stages are in the user's mind. The user normally will already know a lot about the real world represented in the display, and may be able to infer using prior knowledge elements of the interesting structure that have been lost on the route to the display. The visualisation, then, may be of considerably higher entropy than the display itself. From that high-entropy visualisation, the user extracts the low-entropy structure that provides (with luck) the answer to the question. The decreasing and increasing entropies at the different stages of using a visualisation to answer a question as schematized in Figure B-1 (copied from Figure 2-4).

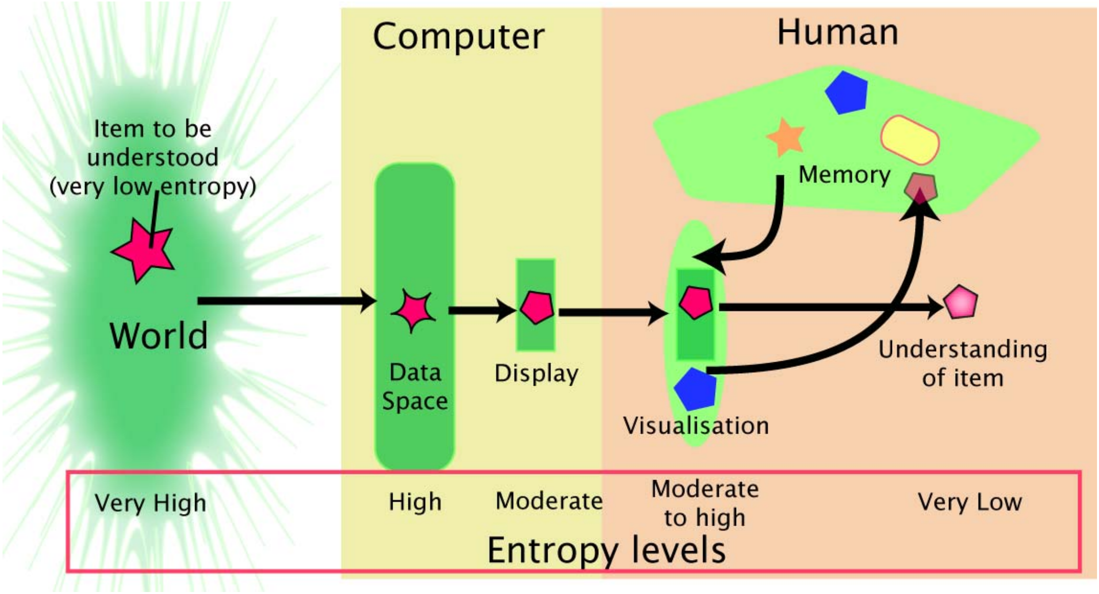


Figure B-1 (copied from Figure 2-4): Schematic Showing Changes of Entropy as the User Obtains a Small Amount of Task-Relevant Information from the Real World, by Way of Sensor Transfer to the Dataspace, Selection and Algorithmic Manipulation to Form a Display, Visualisation Augmented by the User's Prior Knowledge, and Finally Understanding Based on Visualisation.

B.2 WHAT SHOULD BE IN A FRAMEWORK FOR NETWORK VISUALISATION?

In any Framework for visualisation, three areas must be considered:

- What the user might want to visualise;
- What in the data might be available for presentation; and
- What presentation methods might assist the user to visualise the desired information using the data available.

When the area of interest is specialized, as it is in a Framework for Network Visualisation, all three areas must be analyzed, and the results must be compatible. Ideally, the Framework structure should include effective descriptions or taxonomies of the kinds of things different users might want to visualise about networks, lists of what features of data are useful for the multitude of different possibilities, and at least some indications of what kinds of presentation techniques work well with what kinds of data for what tasks. Such an ideal Framework is quite likely to be impossible to achieve in practice, but it is possible to make a start.

A general Framework for visualisation was developed under the predecessor groups of IST-059, and given the name of the VisTG Reference Model. It considered primarily the way the user's needs reflected on the user's interaction with the dataspace through the intermediary of "engines" that executed processes of data selection, manipulation, and preparation for display. The VisTG Reference Model encompassed visualisation of all kinds of data, and therefore contained specialisations for none. It can, nevertheless, provide the basis for the user-side part of the Framework for any specialization, including the visualisation of networks. The VisTG Reference Model is discussed in some detail in [2] and [4], and further elaborated in Annex H. In the present Annex, it will be described only insofar as necessary to illuminate aspects of the Framework.

A second general approach to a framework for visualisation is the RM-Vis Framework developed by TTCP C3I AGVis (now TP2), which is described in Annex G. Its relationship to the VisTG Reference Model is discussed in Annex H.

Before considering what the user might want to visualise about a network, it is worthwhile to consider what about a network might be available to be visualised. This examination inevitably leads into consideration of the nature of networks in the real world, as opposed to the abstract networks that can be analyzed mathematically. To say, however, that real-world networks involve more than the networks of mathematical analysis is not to dismiss those abstract networks as irrelevant in the real world. Indeed, their analysis is often central to understanding a real network, as is discussed in Annex C. It just is not the whole requirement when it comes to visualisation.

We identify at least five different areas of network description that must be considered when creating a Framework for visualising real-world networks:

- **Network Types**:
  - *Point-to-point, broadcast, stigmergic, fuzzy or crisp, striped, partitioned or unitary*
- **Mathematical Relations in Abstract Networks**
  - *Many important properties (usually considered only in crisp point-to-point networks)*
- **Embedding Fields of Real Networks**
  - *Determine and constrain potentialities of a network in its real-world context*
- **Dynamic Properties of Real Networks**
  - *Changes of network structure, as well as the dynamics of traffic over the network*
- **Transformational Properties and Roles of Nodes and Links**
  - *Real nodes may have specific roles, and may output traffic qualitatively different from their input traffic*

Any or all of these areas may be the focus of the user's purpose for visualising some network at some point in a task. A good display will emphasise those aspects that serve the user's purpose most directly.

Looking from the other side of the problem one must ask for what purposes a user might want to visualise something about a network. It is much more difficult to characterise or to create a descriptive taxonomy of purposes than to characterize the features of networks that might be available for visualisation, since the concept of a network pervades so many different areas, including among many others the networks of influence in the genetic processing in a cell, the detection of intrusions in computer networks, the analysis of vulnerabilities in the electricity supply system, the discovery of key personnel in terrorist groups, the interplay of ideas in scientific discovery, the dynamics of planning a military operation, the optimization of routings of supplies for disaster relief, the effects of changes in traffic light timings or in street signage, the effects of gossip and of advertising on the growth of disease epidemics, and so on and so forth. All these intrinsically embody networks and their visualisation, as do many more equally varied applications.

Accordingly, before we attempt to categorize potential user purposes, we examine more closely the nature of networks. Having done so, we will be in a better position to consider possible classes of purpose.

We consider four dimensions of description that might be interesting to a user. From the widest to the narrowest view, they are:

- (Section 2.1) **Network Situation and Context** (Embedding field hierarchy)
  - *Various network contexts (embedding fields and their hierarchies) may be important.*
- (Section 2.2) **Network Structure Properties** (static and dynamic) within its embedding fields
  - *The network itself may be the thing of interest, rather than the traffic over it.*
- (Section 2.3) **Local Properties of Nodes, Links and Sub-Nets** (Drilling down) within the network
  - *The important items may not be the network, but may be found by examination of sub-nets or individual components of the network.*
- (Section 2.4) **Network Traffic Properties** (static and dynamic) processed by nodes and propagated over links
  - *The network traffic, rather than its components or structure, might be important.*

Though these four dimensions of description can all be in play in any one problem, usually one of them is likely to be the most important. We will discuss them in order.

B.2.1 Embedding Fields

Since we do not know of any prior description of the concept of "embedding fields" for networks, they merit an extended discussion. What are they, and why do they matter?

Embedding fields are introduced in Chapter 2 of this report, as follows:

> *Although a graph can exist sui generis, a network exists only in some real-world context. That context gives meaning to the network above and beyond its mathematical properties. To display some context usually helps a user to understand the implications of a display, but at no time can all the context be displayed – it would be the entire universe! The concept of an "embedding field" helps to define the context likely to be useful. […]*

The concept of an "embedding field" was triggered by a pair of hypothesized assertions:

1) *A physical network always has the possibility that a conceptual network lies on top of it. The conceptual network may map homologously onto the physical network if the relationships between nodes are defined as such, but in most cases, the conceptual network involves only sub-sets of the physical network.*
2) *A conceptual network may exist without any underlying physical network.*

Examining these assertions led to the concept of an "embedding field" for a network with or without a physical substrate.

A network in the real world consists of physical or conceptual entities connected by relationships that may be:

- Physically embodied (e.g. roads, wires); or
- Purely conceptual (family tree, social influence, conceptual relationship, etc.).

A network may be *embedded* in a physical or conceptual substrate, but what determines its "embedding field" is the set of contextual attributes for which changes make a difference to the network ***from the viewpoint of the user and for the user's current purpose.*** The embedding field can be thought of as the currently relevant context. Embedding fields are of two kinds, semantic and pragmatic.

In linguistics, three kinds of relationship among words and phrases are normally specified: syntactic, semantic, and pragmatic. Syntactic relationships are among word types; semantic relations involve the normal meanings of the words or phrases, and pragmatic relations are between the text and the external world. To illustrate, consider the following examples:

- The classic sentence "*Colourless green ideas sleep furiously*" is syntactically unexceptional, but is semantic nonsense.
- "*Theodore Roosevelt enjoyed tea with Julius Caesar*" is syntactically and semantically well formed, but is pragmatic nonsense.
- "*The members of IST-059 never met in person*" is well formed syntactically, semantically, and pragmatically, but is factually false. The difference between this and the previous example is that the known properties of Roosevelt and of Caesar preclude the possibility of their having met, and the properties of Caesar and tea make it extremely unlikely that Caesar ever enjoyed tea. In contrast, the properties of the members of IST-059 make it quite feasible that all the meetings could have been done without face to face contact, and it is simply a matter or recorded fact as to whether that happened to be true or false.
- "*Off shopping*" is, in the right context, pragmatically and semantically well formed (as a response to the question "Where are you going?"), but is not syntactically well formed. As we discuss below, interactive situations relax the requirements for well formed syntax not only in language, but also in display.

We can define a similar set of distinctions in network analysis.

In network analysis, graph theory applies to abstract structures of nodes and links, which can be identified as a syntactic level of analysis. Social Network Analysis (SNA) is concerned with relationships such as "works with", "approves of", and the like, which are semantic in nature. One kind of embedding field, such as the TCP-IP network that supports the Web, is also of this nature. Pragmatic analysis is concerned with the relation of the network and network activity to the world outside the network. A second kind of embedding field, such as the landscape on which a road network lies, is of this kind.

In the immediately following descriptions of some possible kinds of embedding fields, we ignore the important fact that the embedding field to be displayed, if any, must be relevant to the user's purpose. We will initially consider only the possibilities for kinds of embedding fields, and will return later to the concept of them as "relevant context".

B.2.1.1 Supporting (Semantic) Embedding Fields

Any network that is more than a topological abstraction exists within a supporting embedding field. Its properties refine and extend those of the field on which it is supported in much the same way that the properties of a software object refine and extend those of the parent object in its inheritance structure. The supporting structure thus provides the semantic context for the network, suggesting and constraining what its properties might be.

It is probably easier to provide a few examples of semantic embedding fields for networks than to define exactly what an embedding field is, although a description using mathematical language might be based on the following: A field is semantically embedded in another field if there exists an injective homomorphism from one to some sub-field (possibly the whole field) of the other (i.e. a mapping from one to the other that preserves at least the structural aspects). In the real world, a network is semantically embedded in an embedding field if it depends on the existence of the embedding field in order to function. Sometimes the validity of this last factor is hard to assess. It is not possible to provide even a rough mathematical statement about pragmatic embedding fields since there is no intrinsic limit to the kinds of real-world relationships that might apply to an arbitrary network.

As an example of a semantic embedding field, a computer network defined by the TCP/IP protocol structure and the capabilities of machines to use those protocols is limited by the physical structure of the computer hardware, the wires or broadcast media of communication, and the operating systems of the computers. These latter provide an embedding field for the TCP/IP protocol network. It cannot work faster than the physical properties of the hardware and the operating systems permit, and it cannot link computers that they do not. However, it extends the properties of the embedding field by providing the computers with a way to identify each other, and to interpret the physical signals so as to permit the computers to exchange messages of arbitrary length and internal structure.

The TCP/IP network itself forms a semantic embedding field for the World Wide Web (the Web). The interconnections of the Web are a sub-set of those available to the TCP/IP network, and the connection speeds of interchanging Web pages are limited by the speeds of message passing over the TCP/IP network. The Web extends the properties of the TCP/IP network in a variety of ways embodied in its own protocols, such as HTTP and FTP.

The Web could not exist without the TCP/IP protocol network, even though one can easily imagine building an equivalent Web based on a completely different set of protocols. As matters stand, the TCP/IP network enables the Web, and is an embedding field for it. Another embedding field for the Web is the network of computers, connected by physical wires or wirelessly, over which the information packets are transmitted. This same physical network is also an embedding field for the TCP/IP protocol network. Any particular network might have a variety of semantic embedding fields, perhaps hierarchically organized, perhaps unrelated to one another.

The TCP/IP network provides an embedding field not only for the World Wide Web, but quite independently for a social network whose nodes are people and whose links are defined by the passage of e-mail messages. The embedding field is the same as for the Web, but the two embedded networks are very different in nature. The Web network is a traffic-free network determined by the links that are coded into the many millions of Web pages, and is well-defined at any moment in time. In principle, one could take a snapshot at some instant and identify all the nodes and links that form the network called “the Web”. One cannot do that with the social network defined by the passage of e-mail messages. It is defined only by integrating the traffic over time. At any one moment, only those packets in transit could be taken as defining a network, and a network so defined would be a very small fragmented one, compared to the network that would be defined by summing all the message senders and recipients over, say, a day or a month.

If a network inherits some properties from its embedding field(s), it follows that one approach to representing network properties is to examine the distinction between the properties of the network and those of the relevant embedding field. Those distinctions represent the information that must be added to the viewer's understanding of the embedding field in order to understand the network, and thus should be an appropriate target for an information-theoretic approach to display.

If the user knows nothing of the embedding field, then the required information is just what is inherent in the network itself. But if the user is familiar with the embedding field, to specify the network with reference to the embedding field may take less information. Furthermore, if there is a display that seems well suited to the embedding field, that same display may well form a good basis for displaying the network.

B.2.1.2 Pragmatic Embedding Field for a Network

An embedding field, whether semantic or pragmatic, is a substrate on which at least the nodes, and usually the links, of a network are defined. An embedding field is likely to be physical, but need not be. A non-physical example is the network of thoughts in a human mind. The physical brain provides the physical mechanism for thinking, but is not the embedding field of the thought network. No thought, and no relationship among thoughts, can be identified with a particular brain location or (as yet) brain activity.

Pragmatic embedding fields can have any dimensionality from zero upward, and the dimensionality of the embedding field can constrain the properties of the network. A network can have more than one pragmatic embedding field in addition to its supporting semantic embedding fields, since its nodes may be amenable to description in a variety of ways and in a variety of contexts.

For a road network, the most obvious pragmatic embedding field is the landscape on which the roads were built. The landscape is not a network, in contrast to the semantic embedding fields for the Web. The landscape is a spatial continuum. The landscape is not the only possible pragmatic embedding field for a road network, however. Socio-political circumstances might be equally important, as for example, might be the whereabouts of Taliban forces in consideration of the road network of Afghanistan at any particular moment.

For the structure of a spiderweb, a semantic embedding field could be the physical web itself (there being no non-web solid material constraining it) whereas a pragmatic embedding field could be the three-dimensional air-filled space around the web (which makes the web quiver when there is a breeze).

For a computer network, a pragmatic embedding field for the hierarchy of semantic embedding fields of the Web could be the physical manifestation of computers and cables over which messages pass, which affects the possibilities of signal interference (though with increasing wireless communication this becomes an inadequate description), or it could include as well all the people who contribute to the traffic flow over the network and those who maintain the physical structures. A semantic embedding field for a network of infection is the set of all people who might conceivably have had the opportunity to become infected, including all those who remained healthy, whereas the cultural and social environment of those people forms a pragmatic embedding field for the infection network as well as for the network of social contacts. And so forth; such examples may suggest the variety of forms that can be taken by embedding fields.

B.2.1.3 Semantic and Pragmatic Embedding Fields Together

The idea of the embedding field, then, is of a system of support or influence that is not itself part of the network, but within or on which the network exists. The landscape is an embedding field for the road system partly because the road could have been constructed to one side or the other of its actual location without affecting the places the road connects, and because if the ground subsides, it affects the road. The air may be considered an embedding field of the spiderweb, because movement in the air affects the behaviour of the web, without affecting the connections.

The individual words of a discourse form an embedding field of a different kind for the linguistically independent syntactic and semantic networks that can form over it. For the linguistic syntactic network that connects the words in a discourse, the embedding field is the string of consecutive words. The set of words is at the same time an embedding field for a quite different network, the linguistic semantic network of associated meanings that connect the words of the string, but in this case the linked words may be far separated in the underlying word string. The syntactic network is described in many textbooks as hierarchic, whereas the semantic network displays many of the characteristics of a scale-free or small-world network (described in Annex C). For these networks, there is no physical manifestation of either of the networks or of the embedding field (the shapes on the page or the sounds in the air are not the words; the words are concepts in the person's head). The linguistic syntactic and semantic networks exist only in the relationships among the words. In this, it is much like the landscape at the time roads are being planned. That landscape is not the physical rocks and soil on which the road will be built, but a concept. In the design phase, the landscape is a conceptual embedding field for a conceptual road network.

The concept of embedding field is recursive, as what is "the network" in one view may be a semantic embedding field of another network. For example, the network of contagious infections has the network of social contacts as an embedding field. For a multilevel example, the TCP/IP protocol software forms one network with the network of physical computers and their links as its embedding field, but the TCP/IP network is itself the embedding field for the World Wide Web, and the Web is the embedding field for innumerable networks of interest among the users of the Web. Each of these networks inherits and augments the properties of its semantic embedding field. In this respect, the relationships of networks with their semantic embedding fields are akin to the inheritance relationships of classes in object-oriented programming.

B.2.1.4 Dimensionality of an Embedding Field

A network that can be represented by a graph including both nodes and links consists conceptually of lines of dimension 1 that connect points of dimension zero. It therefore has a dimension of 1.0 (we need not consider fractal networks at this point, since they rarely apply in practical cases). A broadcast network might be thought to be of higher dimensionality because its links are embodied in the field of, say, radio waves. This would be an incorrect view, since the network links are between the broadcast transmitter and the individual receivers. The broadcast field of radio waves is an element of the pragmatic embedding field of the broadcast network, not of the network itself.

Although a network is ordinarily of dimension 1.0, its pragmatic embedding field, as the broadcast example suggests, may have any dimension from zero (e.g. the words of a text) to at least three (and more if time is considered a dimension, as it would be in a study of, say, the propagation of ideas and culture among generations of politicians). A semantic embedding field, however, is usually another network, which will have a dimension of 1.0.

In what follows, the embedding fields are mainly pragmatic.

B.2.1.5 Embedding Field with Dimension Zero

An embedding field with dimension zero is one in which the nodes may be specified, without links. The field consists of a set of dimensionless points, some or all of which are identified with nodes of the embedded network. Networks defined over a string of text are of this kind. The nodes are identified with some or all of the words of the text, but the links have no representation within the text. They exist only in the syntactic, semantic, and pragmatic structures that are built in the reader's mind.

Social networks might also have a zero-dimensional embedding field, the nodes being individual people, the links being the occasions when person A meets person B. However, a social network might have an embedding field of greater dimension; one supported by telephone communication has the physical telephone network as a semantic embedding field of dimension 1.0. Another possible embedding field for the same social network is the geographic space containing the residences of the people concerned; this pragmatic embedding field is at least two-dimensional.

If the embedding field has dimension zero, the network is necessarily conceptual, with no physical substrate except perhaps for the physical expression of the nodes themselves. Isolated nodes do not a network make.

B.2.1.6 Embedding Field with Dimension 1.0

An embedding field with dimension unity is one in which the nodes can be located, as can at least some of the links. If an embedding field has dimension 1.0, it is quite likely to be a semantic embedding field.

An example is a wire-connected network of computers. Any network defined by traffic among these computers lies on this set of lines. There is no concept of moving a link "sideways" off the wire that conveys the network traffic. The wires may exist in a three-dimensional "real" space, but the network links exist only within the wires, and are identified by their wires, not by the physical locations of the wires. Hence this semantic embedding field is unidimensional.

Infection networks of contagious diseases (those spread by direct contact) have a one-dimensional semantic embedding field, as the links can be identified with the contact events. However, the same infection network can have a three-dimensional pragmatic embedding field, consisting of the space-time in which the infected and susceptible people move, which determines the likelihood of contact events. Which embedding field is important depends on the user's task. Infection networks of other kinds may have embedding fields of other dimensionalities.

B.2.1.7 Embedding Field with Dimension 2.0

If the embedding field has two dimensions, a link can be imagined as being moved "sideways" to a new location in a way that affects the behaviour of the network. Embedding fields of dimension greater than 1.0 are almost always pragmatic. The network of roads on a map of a landscape is of this kind. The location of the road link between two towns on the landscape is fixed and the road (for network purposes) is unidimensional, but the meaning of the network might be subtly (or importantly) changed if the depiction of the road between those two towns had its curves on the map straightened out. In network terms, it would still be the same unidimensional connector of the two towns, but the practical sense of the link would differ. On a paper map, however, there is no concept of the mapped road moving upward or downward off the paper.

The road network may be a network of dimension 1.0, but the physical road system is of dimension 2.0, at least. Roads have width, and if one is concerned with overpasses, a third dimension must be considered. The physical road network constitutes an embedding field for the highly dynamic network of relationships among the vehicles on the road. Its two-dimensional nature becomes obvious when the relationship between two vehicles degenerates into a head-on collision or a side-swipe between vehicles that should have been in different lanes.

Possibly, one might consider the infection networks of diseases carried by insect vectors to have a 2-D embedding field, since it does make a difference where the insects fly across the terrain, but their vertical movements may not matter very much unless infected birds provide a reservoir of the disease.

B.2.1.8 Embedding Field with Dimension 3.0

If the nodes and links can move in any direction in normal space and the practical meaning of the network changes if they do so move, then the pragmatic embedding field has three dimensions. Alternatively, the nodes and/or links may individually be spread over a finite region of the space, in which case the network itself may have a higher dimensionality – but it would no longer be a network representable by a graph. The wires of a wired computer network can move in their 3-space, but such movements do not alter the network in any meaningful way. On the other hand, the network of stray capacitances within a wired circuit, a printed circuit, or an integrated circuit do alter when the position of any wire is moved in any direction. The stray capacitance network therefore has a 3-D embedding field. The same applies to the radiation field from an unshielded wired network.

The links of a wireless network are, for a different reason, in a 3-D embedding field, as they are spread through the space, and that spread is meaningful, since it carries implications for multiple receptions of a message and for the possibilities of interception by unintended recipients. The electromagnetic environment, in this case, provides support to an ill-defined network, and is thus either a semantic or a pragmatic embedding field, depending on which properties are of concern.

Infection networks of air-borne diseases have a 3-dimensional embedding field, as do some pheromone-based networks of interactions among social insects.

B.2.1.9 Stigmergic Embedding Field

A stigmergic system is one in which the environment retains some change consequent on an earlier event, and that change affects subsequent behaviour of elements in that environment. A classic example is provided by the ruts left by vehicles crossing a muddy field. Later vehicles find it easier to follow the same ruts, thus deepening them and making it still harder for following vehicles to leave the track. Another example pertinent to health networks might be provided by the transmission of a cold virus from a sufferer who leaves virus on a door handle to a person who opens the same door some time later. The network traffic in a stigmergic system affects the network structure, and therefore subsequent traffic. The stigmergic embedding field provides the opportunity for feedback loops that pass from traffic to structure and back. It compresses or flattens time.

A network on a stigmergic embedding field is one in which the traffic to or from a node leaves something behind that influences the behaviour of some other node an indeterminate time later. The long-term potentiation of synapses in a network of neurons is of this kind. Giving someone a piece of information that influences the interpretation of later information is also stigmergic. Broadcasting a radio signal that might be picked up by an indeterminate number of receivers is not stigmergic, since the effect of the broadcast on the electromagnetic environment vanishes if the signal is not picked up immediately. Walking across a sandy beach below the high-tide mark is stigmergic, even though the footprints may be washed out by the next high tide, since a person coming along at any time before then would be able to follow the trail. In network terms, there is a link between the node representing the earlier person (the one leaving the cold virus on the doorknob, or the one leaving the trail on the beach) and the later person or persons influenced by the effect of the earlier person on their common environment.

A network on a stigmergic embedding field is necessarily a broadcast network, but the reverse is not true. Stigmergic embedding fields may be semantic, pragmatic, or both.

B.2.1.10 Inheritance Relationships of Semantic Embedding Fields

The concept of an embedding field has much in common with that of inheritance in an object-oriented software system, but also differs from object inheritance in one important way: the properties and capabilities of the embedding structure importantly constrain those of the embedded structure, whereas in object-oriented programming, child classes can augment the capabilities of the parents in arbitrary ways.

Inheritance in an object-oriented software system is unconstrained, in the sense that the child object can have properties that are unconnected with those of the parent, and are not limited by the properties of the parent. An embedding also offers inheritance of properties, but in this case the properties of the embedded object are constrained by those of the embedding, even though they may be of kinds not defined in the embedding.

Consider the example of a packet-switched TCP/IP network embedded on a physical wire network. The property of “packet” is in no way implicit in the voltage variations that are possible on the wire; looking in the other direction, the concept of “voltage” nowhere appears in the properties of a TCP-IP packet. Nevertheless, the existence of the packet is constrained by correlations over time in the values of the voltages. Most importantly, the amount of information transmissible in packets is limited by the bandwidth of possible changes in the voltage. Going up a level, the idea of a “Web Page” is unrelated to anything in the concept of a TCP-IP packet, but everything in a Web page must be transmissible over the medium of TCP-IP packets. This is quite unlike the case of software inheritance; consider for instance “Coloured Polygon”, in which nothing about the edges and vertices that are properties of “Polygon” constrains the concept of colour.

In all other cases of physical embedding, the concept of “information” constrains the possibilities of inheritance from the embedding field to the embedded object. It is this constraint that differentiates embedding from object-oriented inheritance. And it is this constraint that offers possibilities for the application of information-theoretic constructs to the development of effective displays for network representation (Sections 1.1.1.2 and 3.2, Chapter 4 and Annex D).

B.2.1.11 Embedding Field for Display

Although this section is about the properties of networks, it seems appropriate here to note that the concept of embedding fields can be applied also to displays, which can be regarded as existing in an inheritance tree of embedding fields, the root of which is the physical hardware of the display.

For example, consider the embedding of a 3-D display on a 2-D display surface. A point in the 2-D display at a moment in time has five dimensions – five properties: X and Y position within the frame, and Red, Green, and Blue values of colour. A point in the 3-D display has a sixth property not available in the 2-D embedding field, Z position, which one might think to be unconstrained by the properties of the embedding field, but in fact it is constrained.

The existence and nature of the Z-position property is indeed novel and not implicit in the five properties of the embedding field, but the ability to represent the Z position value for a point is completely constrained by those five values. One point in the 3-D space can be seen as having a greater or lesser Z value than another only by virtue of the relationships across the values of multiple points in the 2-D embedding field. A more distant point in the 3-D space may be “fogged” (displayed with less saturation and contrast) than a nearer point, for example, or across time, points in the 3-D field may change in a coordinated fashion related to input provided by the viewer. However the 3-D effect is produced, it is produced by covarying or contravarying the relationships among the RGB values of pixels across their X-Y location values. In an information-theoretic approach, redundancy that creates structure across the 2-D location field of the embedding space is exchanged for information about the value of Z.

B.2.1.12 Relation between Embedding Fields of Networks and of Displays

Representations of the Web are often imposed on a geographic map, as, for example, in Figure B-2b. The real earth geography is a pragmatic embedding field for the network, and the depiction of the geography in the form of a map projection is an embedding field for the display of the network. The network is shown as lines connecting points on the map that represent the physical locations of the hardware computers. Both the embedding field and the network's properties related to the embedding field are shown. The mapping between the embedding field of the network and the embedding field of its representation is trivial.

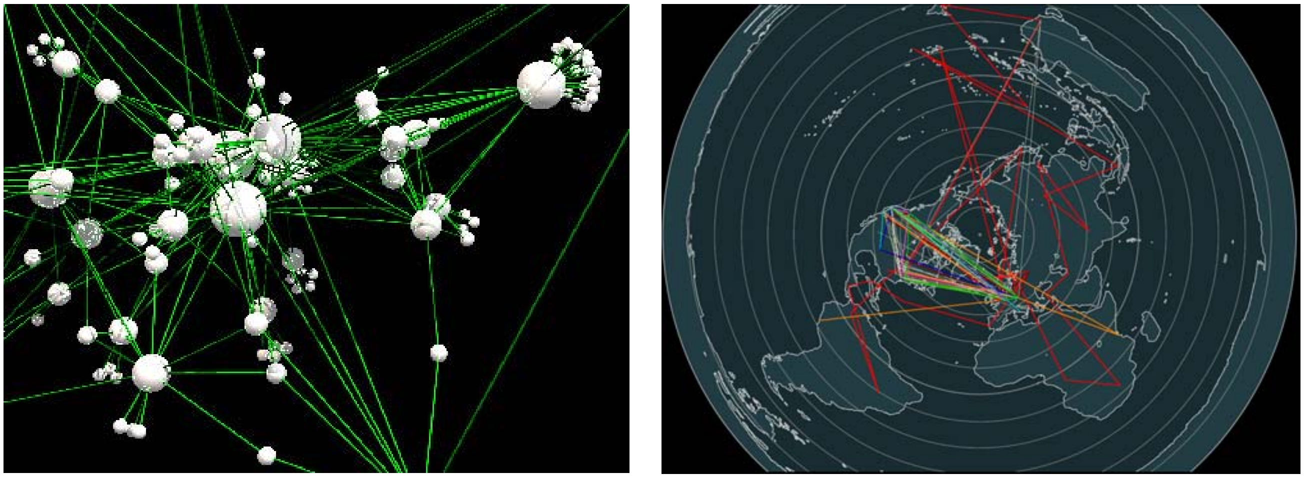


Figure B-2 (reproduced from Chapter 2, Figure 2.6): Two Views on Parts of the World Wide Web. The left picture (Figure B-2a) shows topical relationships, the right one (Figure B-2b) traffic in a geographic context (an embedding field for the network) (Images are from http://www.visualcomplexity.com/vc/, with permission of the respective authors).

In other displays of the Web, or parts of it, geography is of no interest, as in Figure B-2a. The embedding space of the network here is of less concern than its internal structure. But the display has an implicit 3-D embedding space, in which the node representations exist. Distances in the display space represent similarity in the network, and hence the structural measures of the network are represented by the embedding space of the network's displayed representation. To "place" the displayed network in this implicit space, nodes repel one another, but links pull the nodes they connect together, so that nodes (Web pages, in this case) that link to the same node tend to cluster in the display, and if two nodes have a similar population of links to other nodes, they wind up very close to each other despite the mutual repulsion of the node representations.

VITA (Figure B-3) has a similar "artificial gravity" representation, but in the case of VITA, the data space is two-dimensional, with different aspects of the data being laid out on parallel two-dimensional planes in a three-dimensional representational space (a display type sometimes called 2 1/2 D). The user can interactively shift the viewpoint on the 3-D object, which greatly enhances the ability of the user to visualise the shape and structure of the 3-D object. (Other versions of VITA are not necessarily constrained to the 2 1/2 D space.)

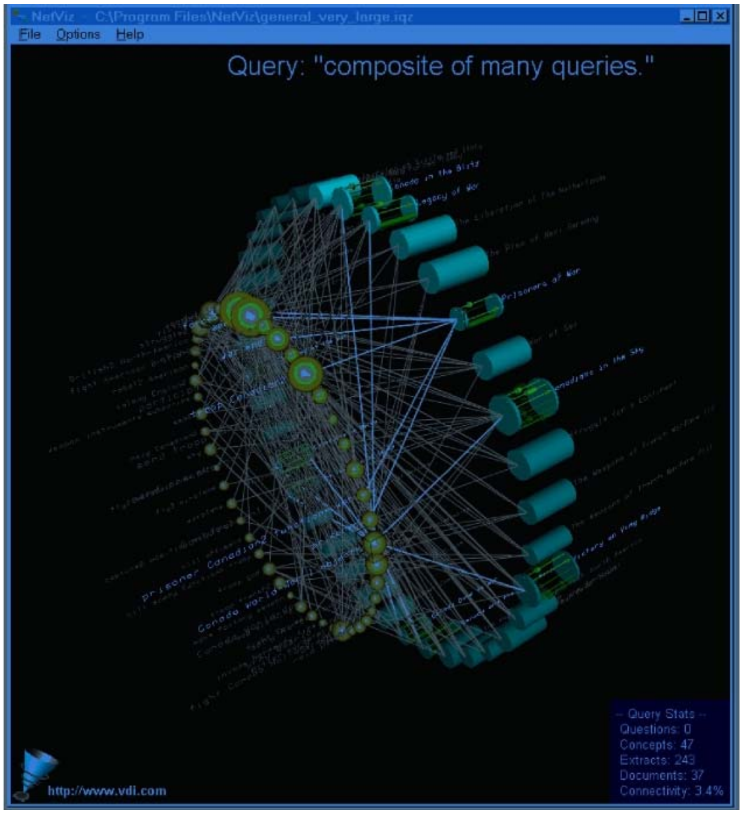


Figure B-3: An Example of a VITA Display, Showing the Three Planes and Some of the Links Connecting Items in Consecutive Planes (Figure reproduced from [2] Figure 7.4).

The planes in the VITA display contain concepts and documents. A third plane, not visible in this example, contains queries to the document database. The network in this case is represented by lines crossing the empty 3-D space between the parallel planes, linking concepts to the documents that contain them. The embedding field of the network consists only of the concepts on one plane and the documents in the other, and thus has zero dimensionality. The conceptual embedding field does not contain the links, any more than the word string discussed above contains the syntactic or semantic relations among its words. The display embedding field, on the other hand, has more than enough dimensionality to accommodate the links.

In the version of VITA illustrated in Figure B-3, the nodes in the network are complex objects, and the user can brush them to obtain information. In the figure several of the nodes represented by cylinders have been opened to show the document title and whereabouts in the document significant information is to be found. This example shows that even when networks are the focus of an investigation, the displays can usefully show information that is intrinsic to a node or link, and even though it may be not an attribute of the network itself.

B.2.1.13 The Network in the Embedding Field

The embedding field is not usually the focus of the user's interest. It is the context for the user's interest in the network. It locates the network with respect to something that the user knows or would benefit from knowing. In Figure B-1, it helps the user to connect what is in the display into the part of the visualisation generated from the user's prior knowledge.

Other than to indicate how novel material relates to what is already known, it is usually not a good idea to display what the user knows already. Display of well known material distracts from appreciation of novel material unless the two are in some way distinguished in the display. This applies equally to the display of temporal variation, where in many cases the static part of the display could usefully be faded, to give the dynamic parts more salience.

Whether to display an embedding field along with the network depends on the user's task. If the task concerns the network pure and simple, then no embedding field should be displayed. If the task involves the context, then at least enough of the embedding field should be displayed to allow the user to be clear about the embedding. How much this is depends on the user's familiarity with the network and its embedding field. Novices will often require more of the embedding field than will experts doing the same task. However, showing more detail is always at the cost of distraction, and it may also make it more difficult for the novice user to distinguish what is shown about the network from what is shown about the embedding field.

Display of an embedding field is likely to be more important for visualisation than for logical analysis, the complementary route to understanding.

B.2.2 Network Structure Properties

There is a vast literature on the structural properties of graphs. We will refer to some of it, but will concentrate more on those aspects of networks that are more difficult to capture as graphs. Social Network Analysis of networks as graphs is considered in much more detail in Annex C.

Network types may be considered from at least two points of view:

- Types of structure or behaviour; and
- Types of real world application.

B.2.2.1 Structural Types

Several different structural types can be identified. Within each there are possibly many sub-types. For example, within the "classic" point-to-point type one can identify such sub-types as "random", "scale-free", "hierarchic", "small-world", and so forth. Such sub-types are not considered in this section.

The basic structural types we identify are:

- **Point-to-Point**
  - The classic network. Nodes are defined and each is or is not linked to each other node. The network may or may not support traffic over its links.
- **Broadcast**
  - A broadcast network must support traffic. A transmitting node cannot know which of many eligible receiving nodes may receive the traffic (e.g. airborne infection, or an over-the-air radio network). Broadcasts may be through a medium in which arbitrary numbers of receivers may exist, or may be over predefined (point-to-point) links on which the transmitting node cannot know whether the potential receiving nodes are active.

Since the concept of "Broadcast" depends on the relationship of one transmitting node to its potential set of receivers, it correctly refers only to a small sub-net consisting of the potential neighbours of the transmitting node. Accordingly, it is quite feasible for a given network to contain sub-nets that are point-to-point as well as sub-nets that are broadcast. Usually, however, nodes that can act as receivers for one transmitting node can also serve as receivers for another, so that the "broadcast" sub-net is more substantial. Such mixed nets may sometimes be better treated and displayed as broadcast, sometimes as point-to-point.

We use the term Broadcast Network in two slightly different senses. The wide sense is defined above. A narrower sense is sometimes used, which is can be made explicit as:

- **Immediate Broadcast Network**
  - Traffic emanating from a node arrives at a potential receiver at some precise time dictated by the environment through which the traffic is broadcast. If not received at that moment, it is never received. This is in contrast to the other kind of Broadcast network, Stigmergic. Often, the term "Broadcast Network" is used loosely as a contrast to a Stigmergic network. The context should make clear which is intended.

- **Stigmergic**
  - "Traffic" is left in the pragmatic embedding field and may be received at an indeterminate later time by an indeterminate number of receivers (e.g. infectious material left on cups or clues left by a criminal for a detective to read; raw material for intelligence analysis is often of this kind).
  - A Stigmergic network is necessarily a type of broadcast network, but colloquially the term "broadcast" is usually taken to mean that the traffic is ephemeral, which means that if it is not received by a node when the opportunity arises, it cannot thereafter be recovered by that node. Colloquially, then, "broadcast" is frequently used in a sense that excludes stigmergic networks. One significant difference between stigmergic and immediate broadcast networks is that the potential set of receiving nodes for an immediate broadcast network could be known to the transmitter, whereas for a stigmergic network this is not true, since some receiving nodes may come into existence long after the initiating node produced its traffic output.

The above types of network may also have the following properties:

- **Fuzzy or Crisp**
  - A "crisp" node or link either exists or it does not, though a crisp link may have a "weight" (discussed below in Section 2.3.1) of any value. A fuzzy node or link is one for which the existence is not well defined. Nodes may be somewhat linked to other nodes (e.g. degree of susceptibility to infection), rather than simply being or not being linked to particular other nodes. The degree of linkage may depend on the user's purpose. Nodes also may be fuzzy. Fuzziness is distinct from probabilistic, though it may sometimes be taken as an aspect of "weight".

- **Multimodal or Coloured**
  - In a multimodal network (sometimes called "coloured"), nodes can be grouped into different classes that have different properties. For example, in a network for battle planning, units may be friendly, hostile, neutral, or unknown (this classification also could be fuzzy, as a so-called "neutral" might be somewhat ill-disposed, but not enough so to be clearly called "hostile", or a "friendly" might not be fully in favour of the objectives of the planner).

- Striped
  - Some multimodal nets may be “striped”. Nodes of type A can be linked only to nodes of type B. For example, a human cannot give malaria to another human, but can give it to a mosquito; a mosquito cannot give malaria to another mosquito, but can give it to a human. A human can move to a location, but a location cannot move to a human. A Broadcast network is ordinarily striped, since some nodes are transmitters, while others are receivers. Striped networks may have any number of different types of node, the node types being distinguished by the types of linkage that are possible. In a striped network, nodes play semantic roles. Nodes of one type perform roles that differ from the roles played by nodes of all other types.

B.2.2.2 Real World (Application Area) Types

Many types of real world networks may be of interest to the user. Some examples:

- Social Networks
- Geographical Networks
- Financial Networks
- Computer/Communications Networks
- Conceptual networks (e.g. the play of ideas or the syntactic links in a text)
- Software structures (e.g. message passing or inheritance structures)
- Networks of influence (e.g. the effects of damage to an electrical sub-station on water or food supply to a town)
- Infection networks
- And many more, limited only by the imagination.

A useful framework must be able to serve all types of networks, including the abstract, the real and the not yet anticipated. As such we must get down to the heart of what it means to be a network. The network in the definition suggested by Working Group 5 of the IST-043 Workshop as “ An array of nodes that exchange ‘stuff’ over links on containers under a certain protocol and following a determined path” has been shown above to be inadequate. Many networks do not “exchange ‘stuff’”. Even if the links do support traffic, there may be no exchange – the traffic may well not be conserved. Moreover, to this definition should be added the words “in some context”.

Nodes may represent people, ideas, towns, banks or bank accounts, computers or other entities, either conceptual or “real”. If there are definable kinds of role played by the entities the nodes represent, the network may be “striped”. The links may be relationships, roads, transfers of money, transfers of information, regions of military dominance or danger, and the like. Link types may determine or be used to discover the roles played by the nodes they connect. The “stuff”, if any, exchanged over these links may be information, diseases, goods, money, and much more. Containers of “stuff” (i.e. traffic) may range from packets (of information), to conversations, activities, trucks and the like. Traffic, if it exists, may be continuous or discrete over time.

Contexts may include supporting networks (such as the social contact network that supports the infection network of a contagious disease) or contexts of higher dimensionality, such as the landscape over which a road network runs. These examples are embedding fields for the networks, but context has a wider range of application. For example, the religio-social environment might be a pragmatic embedding field that could affect the banking network (as happened in Mediaeval Europe, when Christians could not charge interest, but Jews could).

B.2.2.3 Mathematical Aspects of Network Structure

Crisp networks (those in which nodes and links either exist or do not exist) have been well studied over the year, and there is much literature on their mathematical properties. Less work has been done with fuzzy networks, and even in some work that purports to be about fuzzy networks, the so-called fuzziness turns out to be probabilistic and not fuzzy. We will not expand here on the mathematics of crisp point-to-point networks, except for specific cases, in particular to mention one kind of structure that is important in many naturally evolved systems and networks that “just grew” with no overriding design principle. This is the “scale-free” network.

It is possible to take any set of sub-units that interact with one another and define a suitable network to represent such a system [5]. Graph theory has been used to analyze the mathematical properties of many such networks. In Erdos and Renyi’s Random Graph Theory [6][7] it is assumed that each pair of vertices (or nodes) in their network model is connected with probability P, which results in a structure whose number of connections (or edges) per vertex follows a binomial distribution. In most real-world networks, however, the number of edges per vertex decays at much lower rates than in the random graphs. Many such real networks are “scale free”. It is for this reason that they are worth a separate mention.

The term “scale free” implies that whether one views a small segment or sub-net of the whole, or whether one combines small sub-nets into virtual nodes and looks at the bigger picture, the pattern of interconnections is statistically much the same. One cannot tell whether one is looking at a detail or an overview. In other domains of discourse, “scale free” is akin to “fractal”.

A scale free network is characterized by the fact that in any sub-net there are a few highly connected nodes that link the rest of the nodes to the system. This characteristic can be explained by the observation that the number of vertices in naturally evolving real world structures tends to grow constantly, but new vertices tend link to the system through existing vertices that already have large numbers of edges per vertex. This kind of growth and preferential attachment produces a scale-free network.

B.2.2.4 Fuzzy versus Crisp Nodes and Links

Networks are often shown in print and on computer screens as nodes that are represented by dots, connected by links that are shown as lines. Nodes A and B either are or are not connected by a link. This either-or, “yes-no”, dichotomy is called “crisp”. Things are different in the real world that interests us. There, the status of a link or a node may not be crisp. It might be fuzzy. Whether it is crisp or fuzzy, and the membership value of a possible connection in the class “link” if it is fuzzy, may depend on the intentions of the user of the network as much as on the physical structure being represented. Consideration of networks as fuzzy may be an effective way to link the properties of the network to the visualisation needs of the user.

Consider a road network as an example. We can define towns as nodes (and ignore for the moment the fuzziness of the status of hamlets or rural service stops). Also we can define that two towns are linked by road if a traveller can get from one to another along the road network without passing through a third town. To make this concrete, if there are four towns, A, B, C, and D, arranged in a square, and straight roads connect the corner towns A<>D and B<>C, with a crossroad in the middle of the square, then every pair of towns is linked, not just the diagonally opposed ones. If there were a fifth town, E, at the crossroad, then no pair of the original four would be linked, but all four would be linked to E (Figure B-4).

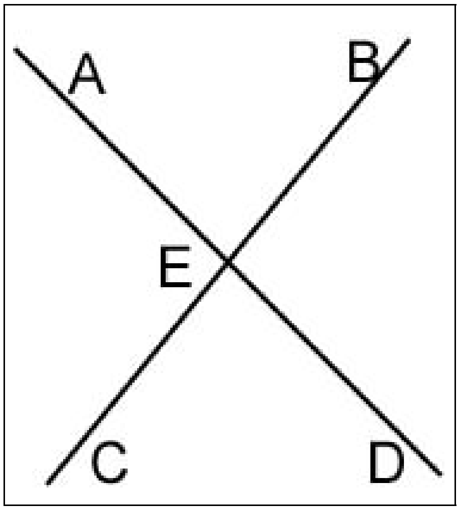


Figure B-4: Schematic Network Showing the Road Links among Four Towns (A, B, C, and D) with a Crossroad at E.

In the foregoing, the role of "the traveller" is overlooked, and it should not be. In network terms, the traveller is the traffic. Whether two towns are linked by particular roads depends very much on the use to which a user wants to put them. Take two extreme examples: (1) a logistics officer who needs to transport large volumes of heavy traffic quickly between A and B, and (2) a hiker who wants to walk pleasantly between A and B. If A and B are linked only by a footpath, there is no link for the logistics officer, but a very good link for the hiker; if, however, they are connected only by a 6-lane expressway, there is no link for the hiker but a good link for the logistics officer. The apparent network that should be represented in any display is different in the two cases.

The interesting situation is the case in which A and B are connected by a two-lane highway. The logistics officer might consider this road be a link, but not a very good one, and so might the hiker. This situation is best represented by asserting that the road has a fuzzy membership in the class "link", and that the membership level depends on the user's intentions for the network. For the hiker, a 6-lane highway has a membership near zero in the class "link", whereas for the logistics officer it has a membership near 1.0. For the footpath, the membership values would be reversed. Both hiker and logistics officer might assign the two-lane highway a membership around 0.6 in the class "link".

This membership function therefore ought to be considered for representation in any display of the network for a particular user at a particular time performing a particular task. It is, in a way, shown on conventional road maps, in that expressways are shown differently from two-lane highways, gravel roads, and footpaths, and scenic roads are marked with green. The hiker may see a marked expressway as having membership zero in the class "link", while the logistics officer sees it as having a membership 1.0.

To provide this variegated symbology is probably about as well as can be done with a display to be viewed passively (Section 4). The map maker does not know whether the user will be driving with intent to go fast, driving with interest in scenery, or will be hiking. But there are more possibilities than those for what the user might want of the road link. For instance, travel time might be of interest to both the hiker and the logistics officer, but their actual time expectations would be very different, and hard to enter on a map without inducing unwanted visual clutter. The situation is different for an interactive display, which can be tailored for the user's needs of the moment.

Returning to the four-towns example of Figure B-4, consider the possibility that a town grows up around the crossroad (town E). If the roads are 6-lane expressways, the growth of the town increases traffic in the nearby section of the expressway, thereby reducing the quality of the link between, say, A and C for the logistics officer. Likewise, if the roads are footpaths, the quality of the link A to C for the hiker might at first be improved as the town comes into being (offering refreshments or accommodations at the newly built pub) and then deteriorate, until as town E grows, it might present a block for the hiker wanting to go from A to C on quiet paths. At that stage, the hiker sees no link between A and C, but there are links between A and E and between C and E. E has become a node for the purposes of the hiker, but perhaps not for the purposes of the logistics officer for whom the town offers no stopping place for his traffic.

The connection between A and C does not lose its membership in the class “link” all at once as a town grows at the crossroad E. It does so smoothly over time, while the memberships of AE and EC (which use the same physical roads) increase their memberships in the class. There is a stage in the development of the five-town array when the connections AE, EC, and AC all have memberships between 0 and 1 in the class “link.” In the definition, therefore, the notion of “passing through a third town” also is fuzzy. The location at E has a fuzzy membership in the class “node”. As Town E grows from a crossroads pub to an industrial powerhouse, its membership in “Node” increases from zero to unity, and indeed, as the town grows, it could split into several nodes, the split also being fuzzy.

Networks being representations of relationships, the same argument can be extended to apply to many different kinds of physical and conceptual networks. Both the relationships and the entities that may be related can be fuzzy, and that fuzziness may well depend on the momentary interests of the user. For example, an intelligence analyst seeking potential terrorists by examining patterns of communication may be better served by a display that shows the likely degree of influence of one person on another than by one that shows crisply whether they communicated.

B.2.2.5 Fuzzy Paths and Cycles

Links in a network can often be concatenated to form a chain or path. In a crisp network, if there is a link between nodes A and B, and another between B and C, then a path exists along the chain between A and C. The situation is less clear if the network is fuzzy. Suppose the connections between A and B, and between B and C have respective memberships of 0.4 and 0.8 in the class “link”; what then is the membership of the route between A and C by way of B in the class “path”? In the network of Figure B-5, is the path from A to C by way of B better or worse than the direct route? It may depend on the user's requirements, as is often the case in the real world.

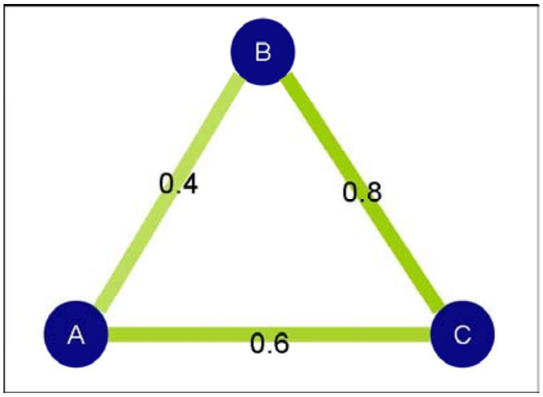


Figure B-5: A Triangular Network of Fuzzy Links.

One possibility for the membership of a connection sequence in the class "path" is the minimum of the memberships of the individual connections in the class "link". In the foregoing example, the sequence A-B-C would have a membership of 0.4 in the class "path". Another possibility would be that the chain would have a membership in "path" equal to the average of the connector memberships in the class "link". The first approach follows the idea that a chain is as strong as its weakest link, and would be suitable if the link membership had to do with its passability for traffic, whereas the second might be more appropriate if the link membership were related to the time it might take traffic to pass, and the question interesting the user were the time to traverse the path.

Now for yet another point of view. If we go back to the definition proposed by WG-2 of the Workshop IST-043, of a network as an array of nodes that exchange "stuff" in containers over links under a certain protocol and following a determined path, then the capacity of the network to support exchange could be a very fitting way to measure its overall value. In social networks, this would measure the capacity for information and/or disease transfer within the network, which directly correlates to "strong" social or relationship links. In computer networks it would mean capacity for information traffic, a measure that correlates to bandwidth and storage capacities. In geographic networks it measures literal traffic and/or volume of cargo, etc. In economic networks, it might correspond to Gross Domestic Product.

Considering the "fuzzy value" of a network in this way would allow stronger memberships in the class "node" and stronger memberships in the set "link" to be more influential in assessing the value of a network or path. This definition of value for a network would necessitate that membership in the class "node" be defined as capacity for storage or processing, and that membership in the class "link" be defined as capacity for transfer. This will take us out of the realm of percentage representations. Instead, we would require a different unit of measure for each type of link in a network: social, geographic, computer, etc. The advantage of this option, however, is that the user can then determine if their preference is for a high-capacity path or a low capacity path, and so forth. This measure would also have some merit in characterizing the overall strength or capacity of the network; which in itself may hold interest to the user.

B.2.2.6 "Striped" or "Alternating" Networks

A pure "Striped" or "alternating" network is one in which there are sets of nodes of different classes, such that for at least one class of node no link can connect nodes of that class to other nodes of the same class. An example might be the network of contacts in a vector-transmitted disease such as malaria. A human can not give malaria to another human, but can give it to a mosquito; a mosquito cannot give malaria to another mosquito, but can give it to a human. The transmission network consists of links from human to mosquito and from mosquito to human, but it contains no links from mosquito to mosquito or from human to human.

In most networks that a user might want to visualise, the nodes and links represent something in the real world. If the network is striped, the node class represents the role played by the entity the node represents. Ordinarily the kind of traffic passed along a link depends on the roles played by the originating and receiving links, so the links also can be segregated into classes. Analysis of the kind of traffic carried over links therefore is one possible method of discovering the striped structure of a network. It is not infallible, however, as the example of malaria shows. The infective agent is the traffic transmitted over the link from human to mosquito and also over the link from mosquito to human, and that agent is the same in both directions.

“Stripiness” is not an all-or-none property of a network. If there are nodes of class A and B, those of class A may simply be more likely to send to nodes of class B than to nodes of their own class, and vice-versa. The degree of stripiness (or the fuzzy membership of the network in the class “striped”) can vary from 1 (no node is linked to another node of its own class) to zero (there is no distributional difference among the nodes of different classes as to the connections of their out-links and in-links).

In a striped network of more than two classes, it is possible that not all nodes have distributional constraints. If a network includes node classes A, B, and X, in which A can send to B or X, B can send to A or X, and X can send to A, B or X, the class X consists of “uncommitted” nodes. Nevertheless, the A-B relationship is striped, nodes of class A and B being prohibited from sending to other nodes of their own class. In such a case, a user might want to visualise the sub-net of A-B connections independently of the entire network, or perhaps might want to investigate the relative density of “X” connections to the other two classes. In real-world applications of network visualisation, especially visualisation of social networks, the roles of the nodes may be as important as the structure of the network.

The concept of “stripiness” applies only to networks that contain cycles. If there are no cycles, then there is no opportunity for “downstream” nodes to connect back to “upstream” nodes. The nodes may have different roles, but those roles cannot be determined or extracted from the network structure. Any one node is a sink for its upstream sub-net and a source for its downstream sub-net, and therefore plays different roles in those two sub-nets, but in an acyclic network that is about all that can be said about role differentiation based on the network structure.

Above, it was mentioned that semantics of a network can be developed from the hierarchy of embedding fields, and that the processing performed in Social Network Analysis is often semantic in nature. Both lead to the concept that a Semantic network is one in which both nodes and links can have a variety of types. The prototypical semantic network represents relationships among words or concepts. For example, one node may be “dog”, another might be “Rover”, and another might be “Paul”. The may be linked in different ways; if the sentence “Paul owns a dog named Rover” is true, then “Paul” --owns--> “Rover” --isa--> “dog”. The two links are of different categories.

The concept of a Semantic network can be seen as an extension of the concept of a Striped network; in a Semantic network only some categories of links are restricted to connecting nodes of one category to nodes of another. Continuing the linguistic example, if “Paul” and “Peter” are nodes of the same category, a link such as “Paul” --knows--> “Peter” would be perfectly normal, but “Paul” --isa--> “Peter” would be disallowed.

If the network contains links that have different properties, it may happen that two nodes A and B are simultaneously connected by several links of different types. Quite possibly, the “stripiness” property pertains only to links of one of these types. For example, mosquitoes can touch humans or other mosquitoes, and humans can touch mosquitoes or other humans. The “touching” links in a network that has both humans and mosquitoes as nodes do not differentiate among them, whereas at the same time the links representing malaria infections define a striped network. The network of complex links may be striped only in respect of some element of the complex link structure.

B.2.2.7 Broadcast and Stigmergic Networks

When a network is drawn as a set of nodes and links, the hidden implication is that traffic on a link that originates at a node will be received by the node at the other end of the link, and only by that node. For two kinds of network, immediate broadcast and stigmergic, this is true only after the fact. In an immediate broadcast or a stigmergic network, the existence of a point-to-point link is determined only retrospectively, by the fact that a node did in fact receive traffic that originated at another node.

In an immediate broadcast network, a node may broadcast its traffic, and any of a large number of other nodes may have the capability of receiving it, but only a (possibly null) sub-set of those potential recipients actually receives the traffic. If the traffic is not received when it is sent, it is not thereafter available to be received. The traffic is transient. The classic example is of a radio transmitter, but examples occur in many fields. A broadcast network is one in which some or all of the links are broadcast.

A stigmergic network depends on a related effect that occurs when a network event alters some facet of the pragmatic embedding field, or environment, of the network, in such a way that the alteration affects the subsequent behaviour or structure of the network. One classic example occurs when a vehicle drives along a muddy road, leaving a rut that induces later traffic to follow the same rut; the earliest published example [8] was of the pheromone trails left by ants foraging for food, trails which guide other ants toward profitable locations and away from unprofitable ones. Another example might be the transmission of disease through infective agents left on surfaces such as door knobs or drinking vessels. This phenomenon is called “Stigmergy”. A stigmergic network is one in which some or all of the nodes act in a stigmergic manner.

The feature common to immediate broadcast and stigmergic networks is that the recipient of the traffic is not known a priori. Nor is it known whether a particular element of traffic emitted by a node will be received at all, or if it is, by how many recipients. The dynamics of broadcast and stigmergic networks are therefore stochastic rather than deterministic.

Broadcast and stigmergic networks differ in the same way that an electrical signal differs from its time integral. In an immediate broadcast network, the traffic is ephemeral, existing only when the “wavefront” of its transmission reaches each potential recipient node, whereas in a stigmergic network, the traffic, once emitted, retains its effect over time and may be received by several other nodes at different later times. Stigmergic traffic may, however, decay or be erased or overlain by later traffic from the same or other nodes, as might happen if a grader came along and removed the ruts on the once-muddy road.

B.2.3 Local Properties of Nodes and Links

Some of the foregoing concerns the properties of individual links and nodes, but with emphasis on the effects of those properties on the larger network. In this section, the emphasis is on the node or link itself.

B.2.3.1 Links

Links may be directed or undirected, elementary or bundled. An elementary link may be directed or undirected, but in either case it represents only one relationship, and if it supports traffic, has only one kind of traffic.

An elementary link can have a variety of properties, such as weight (see below), distance (of a geographical link), flow limit, traffic kind, and so forth, depending on the kind of network, but it cannot be sub-divided into a set of simpler links.

Links may carry traffic or just signify a connection between nodes. A link that carries traffic will inevitably involve delays that might be important to the network dynamics, since the traffic must take some finite time after leaving the originating node before it arrives at the receiving node. The delay time may be fixed (as in the speed-of-light delay between the sending and the receiving of a radio signal) or variable (as in the time a car may take between towns). If the delay is variable, it may be represented by a probability distribution or by a functional process that may take some of its input variables from the embedding field.

Whether the link delay should be displayed depends on the user's task and background knowledge. It may, however, be frequently useful to indicate in the display that a link might have properties that could be viewed on demand.

A bundled (also known as compound or complex) link is a collection of elementary links that connect the same two nodes. For example, person A might at the same time:

- **be** the father of person B;
- **lend** money to B;
- **enjoy** B's company;
- **telephone** B frequently.

A braided link is a particular kind of bundled link in which all the constituent elementary links are of the same kind. If two roads connect A and B, the A-B road link is braided, but if A and B are connected by road and by rail, the A-B transportation link is bundled but not braided.

These attributes have obvious implications for display. A bundled link is a candidate for drilling down to examine the elementary constituent links, whereas an elementary link is not. The user should be able to determine from the display which is the case. As noted above, if the network has bundled links, it may be homogeneous for one class of elementary link, but striped for another class of link within the bundle.

Links in point-to-point networks can have weights or strengths, but what does "weight" or "strength" mean? Several different properties equally might deserve to be called the "strength" or "weight" of a link, some of them at the same time:

- **Utilization** – If the link is of a kind that has traffic, how much traffic is it carrying?
- **Capacity** – How much traffic could the link sustain?
- **Availability** – What is the probability the link will be available for traffic?
- **Number of Braids** – Two nodes may be connected by many parallel elementary links of the same kind (braids) *(e.g. coincidences of the occurrences of names in several documents, where the names are the nodes, and the links their common occurrence in a document)*.
- **Bundling** – If the link is actually a complex of different kinds of connection, weight may refer to the number of different types (*e.g. is a "transportation" link between two towns actually composed of road, rail, and air, connections*).

- **Timing** – A link for which traffic leaving the source node reaches the sink node quickly may be considered to have more weight than one in which the traffic takes a long time between nodes.
- **Distance –** A link between geographically distant nodes may have less weight than one between neighbours.
- **Fuzzy membership** – How much like a link is the connection?
- **Coherence between linked nodes** – (Of a traffic-free link) How tight is the relationship between the connected nodes? (*sibling is tighter than second cousin; "see" is more closely related to "view" than to "grow"*)

A single elementary link may have all of the first seven properties at the same time, and possibly may also have a fuzzy membership in the class "link". In the case of a road link, for example, it is important to a traffic analyst to know how heavily the road is used at a particular time of day (utilization), to a road planner to know its capacity, to a maintenance engineer to know how many lanes (braids) it has, and to a driver planning a route to know how likely the road is to be open (availability), how long it is (distance), and the expected travel time from start to destination (timing). A commander planning a manoeuvre might want to know all of these: How much is this bridge used by the local population, will it carry the weight of my armour, and how likely is it to have been destroyed by the enemy before I use it? Displays need to use more than simply the thickness of a line to show all these different concepts of the "weight" of a link. Add to that the possibility that the information available about any or all of the measures may be uncertain, and the display problem is made appreciably more difficult.

B.2.3.2 Nodes

In a graph, nodes are places where links meet. In real life, nodes often are complex processors, and their out-links may be of quite different character from their in-links. Seldom are the nodes in an interesting network simply places where traffic coming from one link is redistributed to one or more outgoing links. As with links, this potential complexity argues that displayed nodes should indicate whether they have structural information that could be shown at the user's command.

With links, some characteristic properties apply fairly widely, as suggested above. Other properties are hard to categorize, because they span the whole spectrum of possible relationships. When it comes to nodes, characteristic properties are hard to come by; any node with traffic can be seen as a processor accepting inputs, probably asynchronously, and emitting outputs over time, whereas any node without traffic can be seen as an arbitrary network connecting its inputs to its outputs. Accordingly, the properties of nodes are those of arbitrary processors. About the only generic properties that can be asserted are the characteristic dynamical relations between the inputs of a node and its outputs in a network with traffic.

B.2.4 Traffic and Dynamical Effects

Any network of interest to a user is likely to represent something about the real world, and it is likely to be the real world rather than the network itself that interests the user. In the real world "things happen". The static structure of the network is, in those cases, merely a context that determines the implications of the happenings. One could look at the network structure as providing an embedding field for the happenings of interest. It constrains the possible traffic dynamics. As a trivial example, an isolated acyclic network cannot sustain oscillatory dynamics in its traffic, whereas a network with link delays and amplifying nodes is highly likely to have oscillatory traffic dynamics.

Two kinds of “happenings” might interest a user. One kind is concerned with the behaviour of the network traffic, and the other concerns changes in the structure of the network. Changes in network structure affect the possible dynamics of the network traffic, but have their own interest. For example, a user may want to know where there are vulnerabilities in an infrastructure network, and what changes to the network might reduce those vulnerabilities. Conversely, the user might want to know what changes in a network have happened as a consequence of some event, and the effects of those changes on the traffic dynamics.

Traffic consists of anything that passes from one node to another, whether it be conceptual, such as information, or physical, such as vehicles on a road. The existence of traffic implies the existence of a link between the two nodes, as well as some level of temporal dynamics in the network: a node that receives traffic on an in-link has one state before and another state after receiving a unit of traffic. Those states may be indistinguishable, as, for example, the state of a road crossing before and after a car passes, but if the node does any processing, the possibilities for dynamic changes are boundless.

Traffic can be continuous, as is water in a distribution system, or discrete, as in the case of packet transmission in a communication protocol between computers or of cars on a road. Continuous traffic is subject to different dynamical constraints than is discrete traffic, though on a long enough time scale, the individual units of discrete traffic may be sufficiently numerous to allow it to be treated as continuous, in the same way that “continuous” flows of material are actually composed of a large number of discrete molecules.

Not all links in a network carry traffic. In a semantic network, links such as “Fido” “isa” “dog”, in which “isa” links “Fido” and “dog”, carry no traffic. They just exist, as do the links among Web pages, which exist because the form “http://somewebpage” is written in the code of a page, to be read and perhaps used at the discretion of a human user. The traffic-free network of Web links is supported by the traffic-carrying TCP/IP network, one of its embedding fields. The traffic-free network of links written into Web pages is itself an embedding field for another network that does carry traffic over its links — the network defined by events that occur when a user clicks on, or software follows, one of the links in a page. The traffic-carrying TCP/IP network also supports a network of e-mail contacts, and the links in that network carry traffic, the individual e-mail messages.

The existence of traffic implies that the network has dynamic possibilities, but its absence does not imply that the network lacks dynamics. In a traffic-free network, the dynamics are limited to changes in the network structure, such as the rapid and widespread changes in the traffic-free network that is the set of links that define the World Wide Web. To follow the changes in the structure of a network can, in many tasks, be as important as is observation of traffic dynamics in other tasks.

B.2.4.1 Traffic Dynamics

Not all networks have traffic, but for those that do, time and the dynamics of the network are usually important for the user.

Perhaps the task is detection of attempts to attack a defined computer network (the target network). At base, the only way of addressing this problem is to consider the data packets transmitted between the target network and the outer world of the larger protocol network. However, the dangerous packets are dangerous not simply because of their existence and timing (except in the case of a distributed denial of service attack), but because of the effects on a host computer (a node) when their contents are used by protocol-handling software. Accordingly, both the attacks and the defences against such attacks usually involve some degree of drilling down into the structure of the nodes and of the units of traffic, perhaps of a network that is an embedding field for the network being protected.

Traffic in networks can cluster or even jam up if densities are higher than links can handle effectively. Roads constitute networks for which this happens almost daily in most cities. Nodes may represent points at which traffic processing is delayed, and this possibility leads to the proposition that one representable property of a network is whether the links can hold more than one unit of traffic at a time, and whether there are limits on the ability of nodes to emit or to accept traffic. It also leads to the concept of information bandwidth for links. The local implications of this are considered in Section 3.2, but here we consider it in light of its implications for a network as a whole.

The traffic dynamics of a network is the subject matter of the discipline known as “System Dynamics”, and is a major field of enquiry unsuited for development within a document setting out a Framework for network visualisation. Nevertheless, a few points may be worth making.

Acyclic networks have very little that could be considered “dynamics”. They can have only traffic sources, distribution points, and sinks. Certainly it may take time for effects to pass from the root to the leaves of a tree, and the analysis of the distributions of this kind of delay might be useful in some situations. In the absence of events occurring outside the network in question, the end-point, however, is always a static state. More commonly, dynamics is a notion that applies in cyclic networks, in which the effects of the output of at least one node return later to influence its inputs.

A network has intrinsic dynamical properties, meaning that if left to itself without the injection of traffic from outside, it might converge to a stable state, to a repetitive oscillatory state, or to a “strange attractor”, the signature of a chaotic state. However, the structure of a network does not, in itself, always determine which of these fates the network will fulfil. Even a rather simple fixed network, with different initial conditions, can converge to a stable state, an oscillatory state or to a strange attractor. Not only that, but when such a network has more or less converged to its appropriate attractor, the injection of a pulse of external traffic may well move it to one of the other states.

Despite the foregoing caveat, it is ordinarily true that much of the system’s range of dynamical behaviour, and perhaps its actual dynamic behaviour, can be determined from its structure and the properties of its nodes and links.

B.2.4.2 Network Structural Changes Over Time

At any moment in time a network has a certain structure. At a different moment in time it may have a different structure. How should the differences between these two structures be presented? One possibility is to display the unchanged part of a network in whatever way it was displayed, with the changed parts shown more clearly. On the other hand, the interesting questions may relate to the structural properties of the network rather than to its specific links and nodes. Did the elimination of a node, for example, drastically change the diameter of the net, which might affect the transit time of effects from one edge to the other? Did the addition of a link introduce new cycles and possibly novel traffic dynamics? If so, how should this change of character be represented to the user?

Many of the properties of a network can be described by mathematical equations that result in numerical values (Annex C). The time differences or derivatives of these values can be treated in the same way as the derivatives of any vector of scalar quantities. No special problems arise when the scalars represent properties of a network. One could equally well compute partial derivatives, derivatives of one property with respect to another, and so forth. Those time trends or correlative effects can be displayed in any ordinary fashion.

In real-life (as opposed to mathematically abstracted) networks, non-structural attributes might change, such as the mix of traffic on a link, or the nature of a social relationship. It is often true that changes in a network are more interesting to the user than is the basic structure of the network. For some applications changes may be smooth and global, whereas for others, changes may be abrupt, and even catastrophic (in the mathematical sense).

The more difficult problems for presentation may be those relating to local aspects of a network, such as possible changes in the fuzzy membership of the link between nodes A and B, the changing centrality of a particular node, changes in the roles of nodes or the general "stripiness" of the network, or the introduction of new nodes or links. If a network parameter changes, are those changes due to some localized effect that the user might want to examine further, or are they generalized changes affecting the whole network? The answers will affect the presentations that are supposed to help the user visualise what is happening relevant to the task at hand.

B.2.4.3 User-Directed Simulation

The foregoing discussion concerns the presentation of changes that are occurring or have occurred in a network. Often, however, the user would like to visualise what might happen if certain changes were induced in a network. A commander in a peace-keeping mission might consider the possible effects of making an ally or a neutral of a village headman, as opposed to using force to coerce desirable behaviour; a traffic planner might want to visualise the effects of blocking a road or changing the timing of traffic lights. We are talking about "what-if" simulations of network behaviour.

When the user generates the change in the network, the objective for the new display is unlikely to be to show that the change occurred. Rather, it almost certainly will be to show further changes that are consequent on the user-directed change, contrasted to what would happen if the change were not made, or if a different change were to be proposed. The user is in the "**Controlling**" mode of perception (Section 3.1), almost by definition, since the loop is complete from display through action through change in the display, and it is likely that the user will continue to make changes until the result conforms to some preconceived purpose.

Although the user is actively controlling the behaviour of the simulated network, it is usually the case that the simulation is done in order to predict the effects of particular actions on a real-world network. In many situations, if the action were to be taken in the real world, the effects could not be undone, in contrast to the effects of trying different actions in a simulation. The user must assume that the properties of the real-world network are stable over time, at least over a time long compared to the time between performing the simulation and using the results on the real-world network. The user is **Exploring** the dynamical behaviour of the network under changing conditions. The fact that at the same time the user is **Controlling** the simulated network display is analogous to the fact that a 16th century mariner **controls** the route of his ship in order to **explore** the coastline of a new-found land.

When one is Exploring, the display problem is always to help the user to connect the new information with what is already known. When the new information comes from altering the conditions of a simulation, the issue may well be how best to show the differing results of the various interventions, rather than simply to show the results of any single intervention. User-controlled simulations therefore present a display problem that is different in kind from the passive display of changes in network structure over time or the dynamic behaviour of traffic in an existing (possibly simulated) network.

In the context of passive sonar, Wright [9] demonstrated a display technique that showed the probable effects of manipulating one of the parameters on the ability to detect a submarine. The user could change the parameter value, and the resulting 3-D display was made available, along with displays for other tested parameter values. The user could highlight, and bring up on the main window, the display corresponding to any parameter value that have been analysed (Figure B-6).

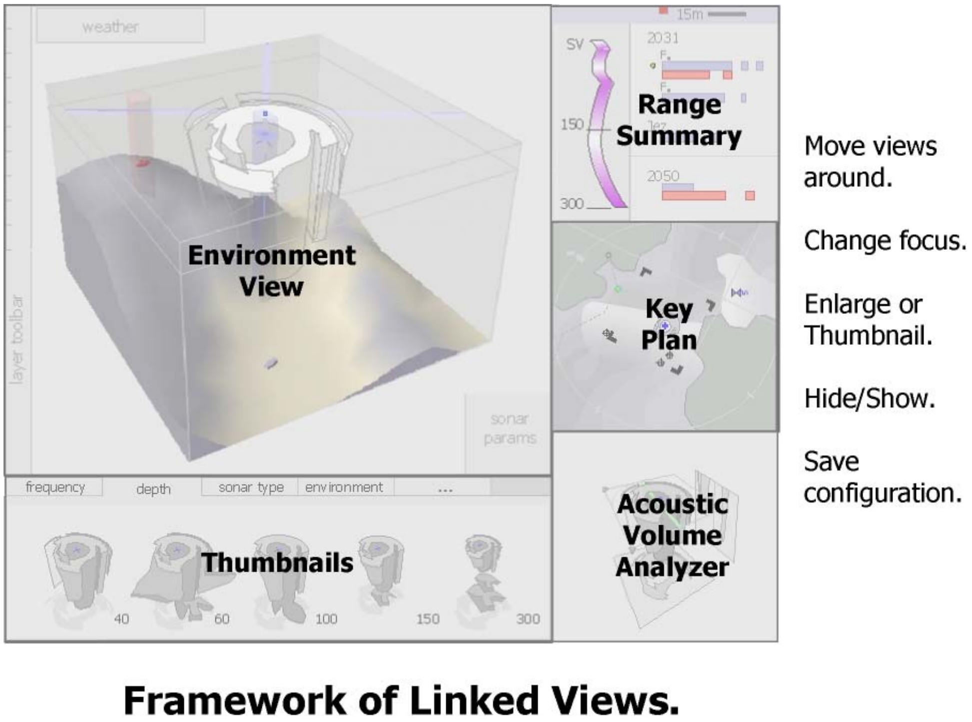


Figure B-6: A View of a Set of Simulations in the Domain of Passive Sonar. The small views at the bottom are for different values of a parameter, any one of which can be shown and manipulated in the main window. (From [7])

The method used by Wright is just one possibility, and it might be harder to use when the changes are dynamic changes in a network. In that case, abstraction of the results into some useful low-dimensional representation might be preferable to an actual display of the network. The visualisation problem is one of attention. It is hard to attend to changes happening simultaneously in several different places. Therefore, if possible, it is better to show the things to be compared in the same general area. For example, in some cases a set of graphs showing the time evolution of the same variable after different manipulations of the network might serve the purpose.

B.2.5 Analytic Abstractions of Crisp Point-to-Point Networks

Most analytic studies of network properties have been concerned with crisp point-to-point networks abstracted from any possible embedding field. In other words, these properties are intrinsic to the networks concerned. Annex C lists a few of them, using social networks as the example type, since social networks have most of the attributes of any crisp point-to-point network.

A social network is a collection of nodes representing members or groups of an underlying population together with ties or links between these nodes denoting binary relationships. As such, a social network is a refinement of a semantic network in which nodes stand for socially significant entities such as:

- People
- Units of action
- Coalition partners
- Departments
- Resources
- Ideas or Skills
- Events
- Nation-states

While the binary relationships indicated by links answer socially significant questions about the nodes they tie:

- Who do you like or respect?
- Transfer of resources
- Authority lines
- Association or affiliation
- Alliance
- Substitution

Annex C presents some of the most useful properties of social networks as abstract semantic networks. It concentrates on mathematical definitions for semantic networks that support measurement of socio-cultural environments of interest.

B.3 PERCEPTUAL ISSUES

The problem of comparing changes that occur simultaneously in different parts of a visual space is only one part of the problem of how to display networks during simulations, and that problem is only one aspect of the general issue of how to display networks. Different kinds of problems become manifest even with one dataset, when the user has different needs. These needs can be characterized at one level by the four modes of perception, recapitulated here from Chapter 2 and from the Final Report of IST-013 [2].

B.3.1 Modes of Perception

As discussed in Chapter 2, perceptions can be categorized according to when and why the perception is used:

- *Monitoring and controlling* use perceptions of changing states of the world in real time, either to ensure that the observed states remain within tolerable limits, or to influence them to approach desired conditions.
- *Searching* also operates in real time. It supports monitoring and controlling when data are lacking in the monitored state, by looking actively for the missing information. The data are used when found.
- *Exploring* is a background activity that does not support real-time monitoring and controlling. Information is acquired about states and structures of the world that are unlikely to change very much by the time that the information may be useful in later real-time monitoring and controlling. When the need arises, prior exploration will have obviated the need for at least some real-time search.

- *Alerting* differs from the other three in that it is a passive process, and in humans likely to be non-conscious and automatic. In computer systems, alerting is likely to be the province of semi-autonomous daemons that monitor the dataspace. The user pre-specifies conditions or states of the dataspace that might suggest a requirement or an opportunity for monitoring or controlling, or that may signal the possible termination of a Search. Humans have evolved comparable internal autonomous alerting systems. An everyday example from human vision is the rapid eye-flick that often follows an unanticipated movement in the visual periphery. The eye-flick allows the person to assess whether the movement signifies something that should be watched, without much distracting from whatever was in focus at the time. Likewise, one readily hears one’s own name in an ongoing conversational hubbub. In computerized systems, alerts can be set so that when an automated process detects a specified pattern in the data, an output is generated that triggers one of the human alerting systems. For example, a portion of the visual display might blink or be shown in an unusual colour, or the sound pattern of an ongoing process might change when one of the daemons has detected the existence of the condition it was set up to notice.

B.3.1.1 Taxonomy of Data Types

Although the four modes of perception determine how the data will be used, the data themselves have a lot to do with how they should best be displayed. A six-dimensional taxonomy of data types was presented in [1], and it can be extended when the data are known to represent a network. First, we present the original taxonomy.

Table B-1: Summary of Data Types

Acquisition	**Streamed**	**Regular**	
		Sporadic	
	Static		
Sources	**Single**		
	Multiple		
Choice	**User-Selected**		
	Externally Imposed		
Identification	**Located**		
	Labelled		
Values	**Analogue**	**Scalar**	
		Vector	
	Categoric (Crisp)	**Symbolic**	Linguistic
			Non-Linguistic
		Non-Symbolic	Linguistic
			Non-Linguistic
	Categoric (Fuzzy)	**Symbolic** (Non-Linguistic)	
		Non-Symbolic (Non-Linguistic)	
Interrelations	**User-Structured**		
	Source-Structured		

Of these dimensions, the first three (Acquisition, Sources, and Choice) probably are the same for networks as for any other data. Data for a network, as for anything else, may be predefined in the dataspace or be incoming on a regular or a sporadic schedule; data may come from one source or many, for a network as for any other kind of material; and for networks as for anything else, it might be the user who chooses what data is available (through, for example, sensor redeployment), or the user might be the passive recipient of whatever data comes to hand.

The final three dimensions, however, may have specific possibilities in the case of networks:

- **Identification**: In the general case, the issue is whether the data objects are identified by the location in a space such as a map or are identified by a label specific to the data element.
  - **Located**: If the data are of a network, location might be relative to at least one of the network's embedding fields, or it might be relative to the rest of the network.
    - If the network data are located relative to an embedding field, is this field semantic or pragmatic. Is it itself a network, is it a zero-dimensional embedding field that supports just the nodes of the network, or is it a spatial extent within which the nodes and links are located?
    - Locating network data relative to the rest of the network means identifying the nodes between which a new link is placed or to which a new node or sub-net is linked. To locate new data in this way implies that the relevant existing elements of the network are already identified.
  - **Labelled**: Labelling means the provision of an identification value to a node or link, or to a sub-net considered as a unit. That a node, link, or sub-net is labelled does not imply that its place in the network is described. For example, a commander may learn that an enemy general has arrived in the area of concern, without learning what responsibilities this general has been given. The general is a node with a name in the network of the enemy order of battle, but his place in it is unknown.

- **Values, Analogue**: Analogue values apply generally to structure measures or to measures local to a single node or link. Both structural and local analogue values may be scalars or vectors, so that dimension of description is not expanded here.
  - **Structural**: Measures such as link density, diameter, minimum cycle path, and so forth, apply to the network or a sub-net. These are ordinarily the result of analysis on data previously available, but nevertheless they are displayable properties of the network; as such, they are data. Typically, structural values are user-selected, single-source, static, though other possibilities exist.
  - **Local**: Local analogue values concern properties of the individual nodes and links of the network. Some of these are intrinsic to their place in the network, such as the numbers of in-links and out-links of a node, or the capacity of a traffic-bearing link. Others are internal to the node or link, and the nature of these data depends entirely on the domain of discourse.

- **Values, Categoric (Crisp or Fuzzy)**: A crisp categoric value is like a label. The thing labelled either is or is not a member of a category. This is contrasted to a fuzzy categoric value. A cat is (crisply) an animal and is (crisply) not a bird or a house. A 190 cm man is (fuzzily) tall and also (fuzzily) of medium height, but (fuzzily) not short. In a multimodal or striped network, the category of a node is a crisp value, but whether an entity is truly a node or another entity a link may be fuzzy. The dimensions of *symbolic* and *linguistic* apply equally to crisp and fuzzy categories.
  - **Symbolic vs. Linguistic**: Symbolic data refer to something, in the way words refer to experiences. Symbols need not be linguistic. On North American roads, a yellow triangular sign is a symbol for caution. In network terms, there is no obvious symbolic element, though the display of a network may use a variety of symbology, such as colour for nodes of different categories.

- **Linguistic vs. Non-Linguistic**: Linguistic data need not be verbal. In [1] the distinction is described as follows:

> *Linguistic data includes more than just words of a natural or a formal language. Any data set that approximately conforms to a known syntax can be described as "linguistic." This includes, say, the structure of the screen display of a personal computer, which has well defined types of elements such as menus, windows that themselves have components such as scroll bars and close boxes, and various other depictions that have properties indicated by their shapes and locations. To be classed as linguistic, the data elements are of a variety of categoric types, each of which has properties that include the influences of elements of one type on those of the same type or another, as an adjective influences its noun, or as a verb mediates the influence of its subject on its object.*

Network data are inherently linguistic in this sense, at least insofar as nodes and links are different categories that never connect to members of their own category. However, attributes of nodes may be either linguistic or non-linguistic. Attributes of links are less likely to be linguistic, though they may be symbolic.

These dimensions of description for data apply to what might be called "atoms" of data, whether they be the values of structural properties or the amount of traffic on a link in a one-hour period. The nature of the data plays a role almost as important as the role of the user's task in suggesting useful kinds of display. In any interesting dataset, however, there are likely to be atoms of many different kinds, some analogue scalar streamed, some static categoric labeled, and so forth. If the dataset is a network that is more interesting than a simple graph, such variety is almost inevitable. Accordingly, even for one given task, it is likely that more than one type of display could be profitably used. We should expect multiwindowed displays to be the norm.

B.3.1.2 Taxonomy of Display Types

The so-called HAT Report [1] described a taxonomy of display types for the display of general information. This display taxonomy has some descriptive dimensions in common with the taxonomy of data types. This taxonomy is reproduced in Table B-2. It may be extended for networks, but the general descriptive dimensions apply whether or not the data refer to networks. We discuss the general case.

Table B-2: Summary of Display Types

Display Timing	**Static**	
	Dynamic	
Data Selection	**User-Selected**	
	Algorithmically Directed	
Data Placement	**Located**	
	Labelled	
Data Values	**Analogue**	Scalar
		Vector
	Categoric	Linguistic
		Non-Linguistic

The first dimension in the display taxonomy is "Display Timing". The issue here is whether the display consists of a static image that holds the entire information to be seen, or whether the information can only be extracted after viewing a changing display over a period of time. When we say “image” and “display”, the same question applies to presentations to non-visual modalities: does the display change informatively over time?

The second dimension requires perhaps a little more explanation. In a sense, the data on a visual display screen or that generates a sound pattern is necessarily algorithmically generated. The descriptive dimension goes a little deeper. It deals with the Controller in the MVC way of looking at the process. The dimension was described in [1] as follows:

> *In a large dataset, only a small portion can be viewed at any one time. That portion might be a few elements of the original data, but more probably it is a distillation of the data – perhaps a set of a few dozen weekly averages to represent a few billion network events, or a representation of an area on a map as “forested” in place of a depiction of the photographic representation of every tree. The data-selection issue is how this reduction of the dataset into a viewable sub-set is accomplished. Is it done by a predetermined algorithm or is it done in response to moment-by-moment choices on the part of the user? Can the user navigate the viewpoint through the possible abstractions of the dataspace?*

In other words, are we dealing with an interactive display in which part of the interaction is the user’s choice of what data is selected for display?

The third dimensions is Data Placement, whether individual data elements are placed on the screen in locations that correspond to some attribute of the element, as would be, for example, the boundaries on a topographic map, or is it placed in a place that depends on the element’s identity (as would be, for example, the link between node i and node j in the matrix representation of a network.

Finally, the fourth dimension is that of Data Value. In this dimension, although the possibilities have the same labels as in the case of the Data Type taxonomy, the implications are different. Now we are dealing with what is displayed on a screen or is represented in sound or some other modality. Analogue data values refer to things like brightness (scalar) and colour (vector) or x-y location (vector) that can be shown on a screen, or in auditory presentations to loudness and pitch (scalars), timbre (vector), and waveform envelope (vector). Categoric data values do not change with the brightness of colour in which they are displayed. A word is that word, no matter what its brightness. Data displayed as words is inevitably categoric and is likely to be linguistic.

As we discussed in respect of the data taxonomy, the term “linguistic” here means obeying some kind of syntax. Words may fail this criterion, and visual display representations may have their own syntax. For example, on a windowed computer screen, the window has elements such as its frame, a menubar, or a scrollbar that are non-verbal, and that have roles indicated by combinations of their shape or colour and their location with respect to the window frame. These elements are categoric linguistic data in the display, and the user interprets than as such.

The data taxonomy and the display taxonomy would be simply amusing exercises, were it not for the fact that data of some types map quite readily onto the display taxonomy, thereby suggesting the general nature of a display that would be effective for those data. Some such mappings are described in [1]. We have not considered their extension to the representation of networks, but it is intended that this be done as part of the further development of the Framework.

B.3.2 Information-Theoretic Issues

The concept of information theory has become very much muddled since Shannon’s original exposition in [10]. Some of the muddle may be due to misunderstanding about what Shannon wrote, or to a view of the nature of probability that is inconsistent with Shannon's formulation. In considering information-theoretic issues, we go back to Shannon's original concept, which we sketch here to clarify some possible ambiguities.

Shannon was interested in communication, and therefore concentrated on the communication channel. His question was how well and how rapidly could information be transmitted from a sender to a receiver over a channel of certain defined characteristics. To address this problem, he identified before and after states in the receiver. Before the transmission, how much did the recipient know about what the originator would transmit, and after the transmission, how much did the recipient know about what had been transmitted? He identified "how much the recipient does not know" as "Uncertainty". The difference between the before and after measures was identified with "information". Hence, contrary to what is often said, since the recipient's uncertainty is always about something, so "information" in the Shannon sense is always about something. When we are talking about displays, the question is what the user knows about the thing displayed before and after looking at the display.

"Uncertainty", according to Shannon, is a property of the recipient of the transmission. A change in the recipient's uncertainty about what the originator would or did transmit is the information conveyed. The measure is subjective, not a measurable absolute quantity. The probabilities involved in the mathematical description are subjective probabilities assigned by the recipient, not absolute values that can be determined from physical considerations, although it is quite possible for the recipient to use physical considerations in developing the probabilities in question.

When we deal with the "information content" of something, we presume some kind of observer, with some kind of prior knowledge, who, by observing the thing, could alter that knowledge by a measurable amount. Quite often in discussions of information and uncertainty, the prior knowledge of the observer, and even the necessity of presuming an observer, is forgotten. Similarly, the fact is often forgotten that all probabilities are conditional on some defined state. A coin has a probability 0.5 of falling heads – or does it? It does not if it is an unbalanced coin or if the coin tosser is skilled in causing it to turn in the air a precise number of times. So the probability that a coin will fall heads with probability 0.5 is conditional on the toss being fair and the coin being balanced. All probabilities are conditional on some prior state or process, even if the precondition is not specified.

B.3.2.1 Information-Theoretic Issues of the Data

B.3.2.1.1 Uncertainties and Information are Observer-Dependent

The need to specify the precondition for an estimate of probability becomes clear when we ask about the "information content" of a network, or of new data about a known network. What does it mean to talk about the information content of a node, a link, a sub-net? It means nothing unless we specify an observer with some kind of defined prior information, and some properties that the observer is noting about the network.

Imagine a simple graph consisting only of links that connect nodes pairwise. No matter what interests the observer, the only data that can be obtained from the network is whether there is a link connecting node A to node B. Or is it? If the observer has not looked at the network, the number of nodes is unknown. How much information would the observer get from determining the number of nodes? If before looking at the network the observer had literally no preconception, all possible numbers being equally likely, and after the observation knew how many nodes there are, the information gained would have been literally infinite. Infinite quantities are not very useful, and it unlikely that in any real case the user would have absolutely no preconception about the number of nodes. More probably, the user would be able to guess quite accurately whether the number of nodes was closer to ten than to a trillion. So, the amount of information available from a specification of the exact number of nodes depends on the observer's prior distribution of probabilities for the number, and the probabilities in the distribution are conditioned on the observer's history of observing similar networks.

Usually, we do not worry about the prior knowledge of the observer, and compute what appear to be well defined probabilities for the attributes to be observed. What is the expected probability distribution of the number of links connecting to an arbitrary node? To answer that, one must imagine a process for assigning links between nodes. That process is the precondition from which probabilities make sense. Suppose that it is known that there are N nodes and L links, and that not all node pairs are connected so that $L < N*(N-1)/2$. Then a random process could be ascribed such that links are individually "dropped" onto the network so that each new link connects two nodes at random, the drop being tried again if those two links were already connected.

This process defines a probability distribution for the number of links connecting to any arbitrary node, and it is easy to assert that this same probability distribution is the prior probability distribution of the observer. Given that, the information gained by observing the actual number of links to a given node can be computed: it is the initial uncertainty of the prior probability distribution, since after the observation there is no residual uncertainty.

Experts and novices differ both in the amount of prior knowledge they bring to an observation and also in the uncertainty that remains after an observation. Before the observation, the expert has less uncertainty than the novice about the property to be observed, so if they both end up with the same uncertainty, the expert has gained less from the observation than has the novice. However, especially in complex situations, it is often only the expert who is able to glean much information from an observation. The novice may well be as uncertain after the observation as before, whereas the expert may have seen some structure in the data that clarifies everything.

As an example, consider a layout of pieces on a chessboard, a layout that has occurred during normal play. A grandmaster will see this layout as a unit, and is likely to be able to replace every piece on the board after a single glance, whereas someone who knows nothing of chess will be lucky to be able to replace more than three or four of the pieces accurately. In this situation, the expert may well have gained more information from the observation than did the novice, because before the observation both knew only that up to 32 pieces would be distributed over the 64 squares, whereas after the observation the expert knew where every piece belonged, whereas the novice knew where only a few belonged.

Displays suited to use by experts are very likely to differ from displays appropriate for subject-area novices. What is intolerable complexity to a novice may be necessary structural detail to the expert. A display that looks good to a designer may not satisfy the expert, simply because it is too uncluttered to allow effective visualisation. A novice may *analyse* using a "clean' display where the expert *visualises* using a complex one.

B.3.2.1.2 Structural Information Measures

Let us imagine the process suggested above, which gives equal probability to any link between nodes. Now we observe the distribution of the number of links connected to the neighbour of a node. This particular process would say that the number of links to neighbour nodes is unrelated to the number of neighbours of the focal node. Looking at all the nodes with only one neighbour, the distribution of the number of links to that neighbour should be the same as the distribution of the number of links to the neighbours of well-connected nodes. In real networks this may not be the case. Well-connected nodes might preferentially connect to singly-connected ones in a hub-and-periphery arrangement, or they might preferentially connect to each other, to form cliques and separable sub-nets. In either case, the hypothetical observer's uncertainty is reduced by observing whether a node is well-connected or otherwise conditioned on the connectivity of its neighbours – the network is structured, and that structure is a candidate for display.

When we consider networks that are more than just graphs, such as networks with their embedding fields, networks with processor nodes, links that are bundles of different kinds of connection between the same two nodes, and so forth, there is much more scope for information-theoretic measures to become important. Always, the user's prior knowledge and current interests will determine what should best be displayed. Is the user controlling? Then the display should be focused on the aspect of the network that is being controlled, using anything else only to maintain context. Is the user Exploring? Then the display should allow the user to shift focus in unpredictable ways. Let us consider a few informational aspects of the network structure in this larger context.

One or more of the embedding fields of a network is usually quite important in setting the context. The user may well know things about an embedding field that constrain the probabilities associated with the network structure. For example, if the embedding field is a landscape, it constrains the field of view from any specified location. If the network is located within such a landscape and nodes are preferentially located with good (or perhaps with poor) fields of view, then the user's prior knowledge of that aspect of the landscape reduces the requirement to display the information for the network. The user's understanding of the embedding field can often reduce the need to display inherited properties of the network.

B.3.2.2 Information-Theoretic Issues of Display

After Shannon presented his seminal work [10] psychologists attempted to use information theory to account for a wide variety of perceptual phenomena. When their analyses turned out not to work in various situations, the initial enthusiasm turned to disparagement, and information theory was almost completely abandoned as a tool. It still has value when applied appropriately, as we hope we do here, and as we think Smestad [11] did in setting forth thirteen principles for creating effective displays and linkages among related displays.

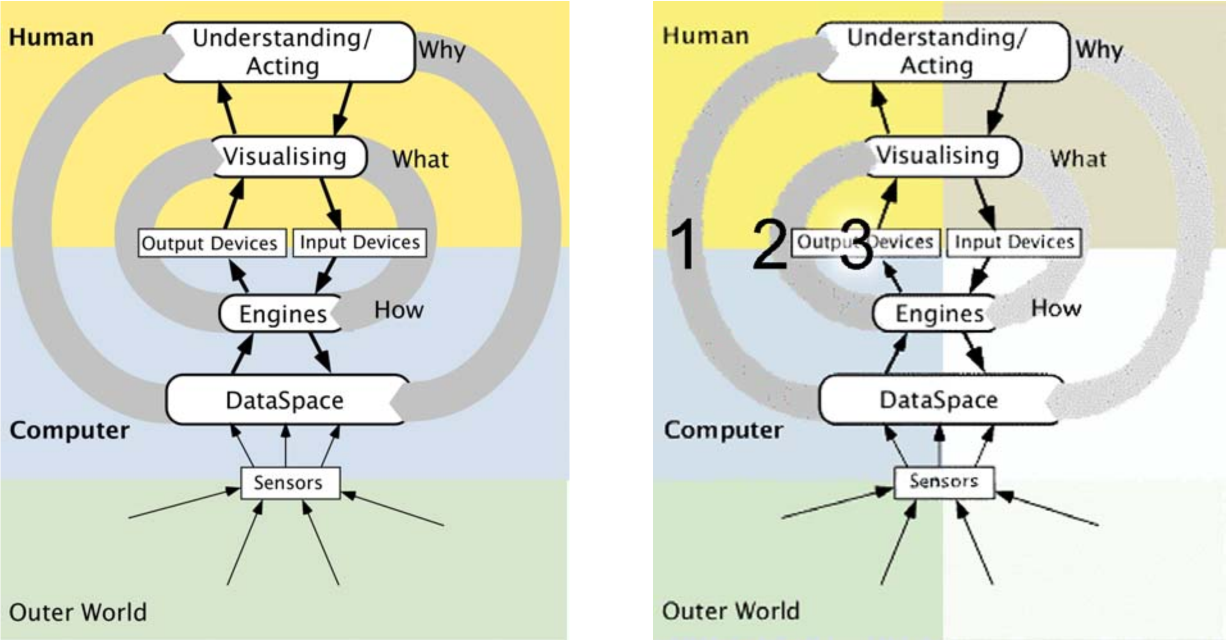


Figure B-7: (a, left) The VisTG Reference Model; (b, right) Three Information Channels that Should be Considered in an Information-Theoretic Analysis of Displays. Channel 1 depends on Channel 2, which in turn depends on the physical Channel 3.

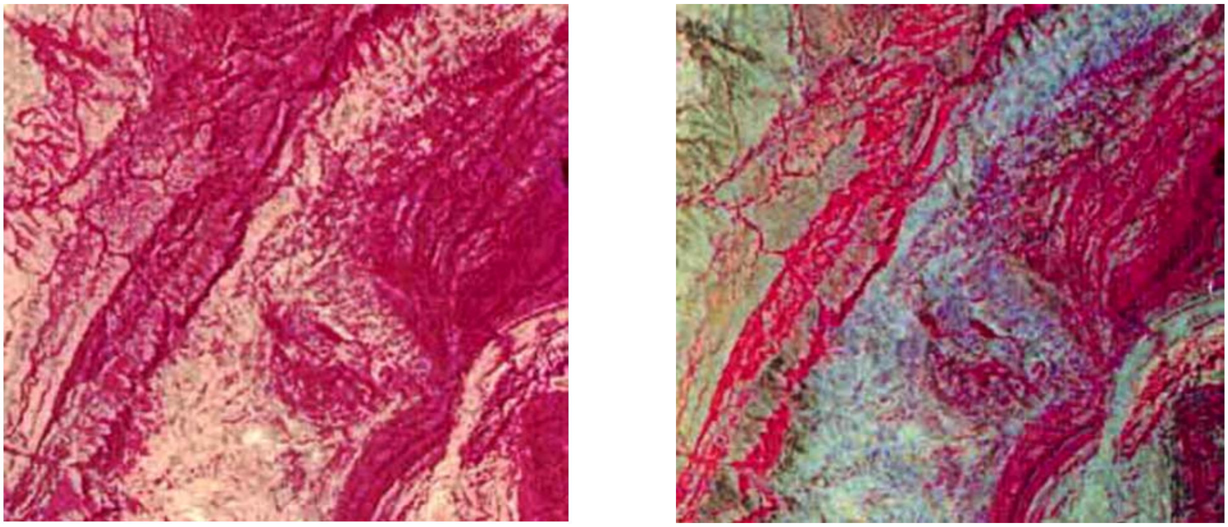
The display is the channel between the dataspace and the user, in Shannon's sense. However, when a person looks at a display, there are several levels of abstraction, each of which can be considered as an information channel for the next level. Consider, for example the left side of the VisTG Reference Model, as shown in Figure B-7b.

The VisTG Reference Model identifies loops at three levels of abstraction, the most abstract being the one between the dataspace and the user's understanding and acting on its implications, the next being between the user's visualisation and the engines that support it (though in practice this loop is rather complex, as discussed in Annex G), and one between the display, the user's eye, the user's muscles, the input devices and the display. Each of these loops has two halves, an output half from the computer to the user, and an input half through which the user influences the computer. We consider here only the output half.

The output halves of the three major loops, from the computer to the human, are labelled 1, 2, and 3 in Figure B-7. We can consider these half-loops to be the communication channels. Channel 3 is a physical channel. It is ignored in Figure B-1. Channel 2 is virtual, its physical implementation being through Channel 3, Engine software in the computer, and some mental processing in the human. In Figure B-1 it is shown as a single link between display and visualisation, within which the display devices and human sensors are subsumed. Likewise, virtual Channel 1 is implemented through Engine functions in the computer, virtual Channel 2, and further mental operations in the human. In Figure B-1 it is the entire length of the communication chain, or at least the part between dataspace and understanding. Each level has its own uncertainties and communication bandwidth. We consider them in turn.

B.3.2.2.1 Intrinsic Limits of the Visual System

At the display device level, Channel 3 in Figure B-7b, the user's prior uncertainty is of precisely which pixels have what colour. After a viewing event, this uncertainty is reduced, and the information conveyed is the difference between these two uncertainties. The uncertainty is not reduced to zero, for at least two reasons. One is that the colour channels of vision have widely different informational channel capacities. Variations in blue-yellow contrast or, equivalently, in blue intensity have a relatively low intensity resolution and spatial bandwidth, meaning that little information comes from blue-yellow variation except across relatively large regions of space. Display designers have learned never to rely on blue to display detail, especially on a dark background. The spatial bandwidth and intensity resolution of red-green contrast is considerably better, but still worse than that of overall brightness, which has the highest spatial bandwidth and resolution of the three visual channels. This fundamental limitation is the reason that the two images of Figure B-8, which technically have the same information content, provide such a different amount of terrain information to the viewer.



a) b)

Figure B-8 (Reproduced from [1] Figure 2.3): A Multispectral Satellite Image of an Area of the Canadian Arctic in Summer – (a) As normally displayed in "false colour," using one sensor channel as red, one as green, and one as blue; (b) By displaying the first three principal components of the spectral variation as, respectively, brightness, red-green contrast, and blue-yellow contrast. Several terrain differences that are invisible in Figure B-8a, are evident in Figure B-8b, even though both images display essentially the same data. (Images produced in 1976 by M.M. Taylor, then at DCIEM, Toronto)

In Figure B-8a, the first three sensor bands are shown as red, green, and blue respectively (a fourth sensor is highly correlated with these three and contributes little). In Figure B-8b the first principal component of variation across the sensor bands, which potentially conveys the most information, is presented as intensity variation, the second as red-green contrast, and the third, which conveys relatively little information, as blue-yellow contrast. The variations in the two pictures are technically almost the same, but the vectors representing red, green, and blue have been rotated from the "natural" alignment of longer-wavelength sensor equals red into a different space in which variations in the vector of sensor values have been rotated into the directions that best match the information carrying capacities of the human visual system (at least for people who are not colour-blind).

B.3.2.2.2 Redundancy and Residual Uncertainty

The second limitation on the reduction of uncertainty from viewing a display comes from the fact that very large numbers of pixel arrays are indistinguishable by any human observer. One "random" scatter of coloured pixels, or even of substantial patches, is hard to distinguish from another, even when they are shown side-by-side. When a number of, say, black pixels form a straight line among light-coloured ones, it is easy to tell if one of the black ones is displaced, but if an equal number of black pixels is scattered around among the light ones, to see a displacement is very hard.

This advantage of well recognized patterns was taken by Garner and his students [11] to define "good form". In an experiment, subjects were asked to sort patterns of marks on a square grid into sets that seemed "the same", and separately to judge the degree to which the patterns represented "good form". The patterns that belonged to small sets after the sorting were also those judged to have good form: straight lines, X patterns, and so forth. Symmetry is also important. The essential point is that because they had fewer partners considered to be like them, the patterns having good form conveyed more information about the display than did patterns that were not good forms.

Extrapolating this to the display of complex data suggests that the actual information that can be received from an image on any specific screen is far less than the simple distributions of possible pixel values would suggest. Only the discriminable sets of patterns on the screen contribute to the information capacity of that channel. In Section 3.2.1.1 the example of the different information gathered by a novice or a grandmaster from a glance at a chessboard provides another illustration of this principle. To the novice, one pattern of chess pieces is much like another, whereas to the grandmaster, patterns reached in a real game are clearly different from one another. They have "good form" whereas most random arrangements of pieces do not.

The information capacity of Channel 3 is constrained by the number of discriminably different sets of patterns under the viewing conditions of the user's task, not by the number that could be discriminated under optimal viewing conditions for making the comparison. The actual number is unimportant, and will be very different under different conditions. It is important, however, to note that there is an extremely large difference between the information technically available from variations in pixel values and the information that a human could derive from those variations. It is this difference, technically called redundancy, that allows for 3-D representations on a 2-D screen, by taking advantage of correlations of pixel values across regions of the screen separately from the correlations that lead to the "good form" discriminable patterns.

To have a feel for the informational differences that depend on good form, consider as a simple analogy a series of letter strings. If you were allowed only a quick glance in each case at String 1 and at another time at String 2, would you notice that they were different?

- **Random Letters**
  - String 1: wmci slabcvuql skversmt 8c andsi4 jdolsjfgn
  - String 2: wmcl atabcyogl eludnvjs; 8c sntxi4 ydolsjtgn
- **Pseudo-Syllabic**
  - String 1: thim arbustidrate skatimol ad indacol ifamiston
  - String 2: thjm arbustadrote sketimal ad irdacol ifamaston

- **Words**
  - String 1: them identical student to escalator validate
  - String 2: them intensify strident to excalibur valorate
- **Sentence**
  - String 1: they produced representations to demonstrate
  - String 2: they produced dispensations to remonstrate

For most people, as the nature of the strings became more and more like normal English it would be successively easier to tell that String 2 was not the same as an earlier presented String 1. In the last case, even though both strings are proper sentences using similar words, they mean different things. The letter patterns are less different than they are in the "Random letters" example, but the discrimination is easier. The "Sentence" strings are more nearly "good form" than are the "Words" strings, for an English speaker (though perhaps not for a Chinese monolingual). Likewise, in a display, discrimination is easier when what is displayed makes sense to the user.

At some level, all the strings make sense, other than the "Random letters". "Skatimol" has at least the basic structure of an English word, and one can easily pronounce it. The change to "sketimol" might be detected by a careful reader, probably more easily than the change from "slabcvuql" to "atabcyogl", though not as easily as the change between "demonstrate" and "remonstrate".

Returning to the representation in Figure B-7b, the "Random letter" strings may be taken as analogous to what could technically be displayed on a screen. Anything goes. However, informationally, patterns analogous to those strings can be substituted quite freely without making a difference to the human viewer, and therefore convey little or no information. The human could, however, discriminate a bit between displays analogous to the pseudo-syllables, which to the human visual system might mean colour patches, lines, and so forth.

Figure B-9 shows another way of looking at the relationships of the different levels of abstraction in perception and in the world perceived. The early visual system, at the bottom of Figure B-9, distinguishes patterns and perhaps objects. This is at the level of Channel 3 in Figure B-7b.

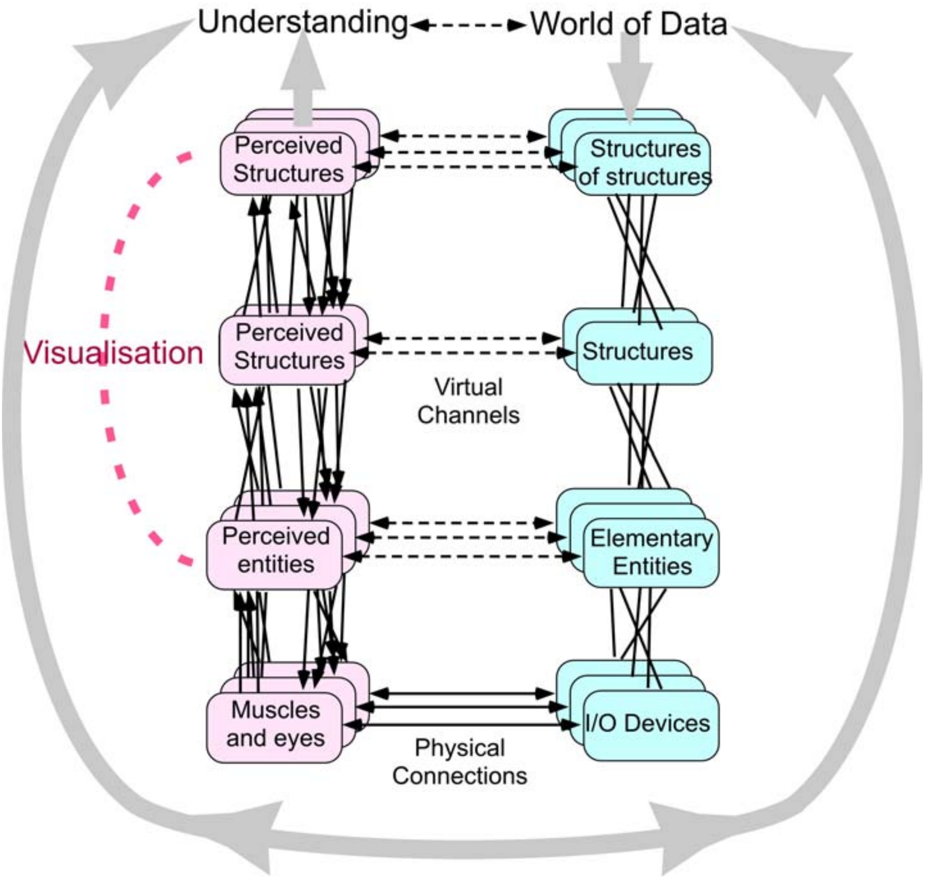


Figure B-9: Schematic Succession of Developing Information Regularities in Visualisation and Understanding. In the context of the VisTG Reference Model Channels of Figure B-7, the middle three levels are all incorporated in Channel 2, the Visualisation to Engine loop.

As the VisTG Reference Model is conceived, the Visualisation system makes sense of the patterns and objects, and the display Engines (Views in the MVC model) create patterns of which the visualisation system can make sense. At this level, however, they may make sense only in that the pictures cohere, by analogy with the Words in the string comparisons above, and at a higher level (still within Channel 2), by analogy with the sentences.

Finally, just as sentences must cohere in a text to make sense in dealing with a topic, so the different visualisations produced as a result of the operation of Channel 2 must cohere to make sense in the context of the user's task and to give the user the impression of understanding the implications of the data in the database, completing Channel 1.

The communication channel for understanding the dataspace is Channel 1, the outer loop of the VisTG Reference Model, which is represented at the top of Figure B-9. Channel 1 is a virtual channel, implemented by the data selection and manipulation engines, Channel 2, and the human visualisation processes. Channel 2, in its turn, is a virtual communication link shown as taking three levels in Figure B-9; it is implemented by the display engines, Channel 3, and human vision (assuming the display is in fact visual). Channel 3 consists of the display hardware and associated software, and human sensory and perceptual processes.

At each level of this cascade, the supporting channel has a higher technical information bandwidth than the channel it supports, the bandwidth at the top level being usually rather low, as suggested in Figure B-1, where the entire end-to-end channel is represented as the top level of Figure B-9. The implication of this is twofold: firstly, different "messages" can convey the same information, and secondly, interpretation in the higher-level channel can be eased by the use of redundancy in the supporting channel.

In Annex D, Bjørke demonstrates how it is possible to compute approximately the informational constraints on Channel 3, and uses the information to produce displays that allow for "informational zoom" as the scale of a display changes interactively (or passively). Bjørke's technique has been used to change automatically the display of road networks as a function of the scale of a displayed map. In this, in addition to the constraints of Channel 3, he uses some of the Channel 2 constraints, such as that the process should not cut the road network into disconnected sub-nets. The appreciation of the connectivity of the network is a function performed in the Channel 2 loop.

B.3.2.3 Information, Entropy, and Modes of Perception

The four modes of perception have different implications for the information rates and display entropies. The object of Exploring, for example, is that the user should build in memory as complete a description of the network as will be useful for later tasks. The information rate might be low, but the eventual structure developed in memory may have a large entropy. Referring to Figure B-1, the display entropy for Exploration should probably be high, allowing the user to explore the display mentally rather than by overt interaction, which requires the diversion of attention from the task to the navigation. Exploring is associated with the channel between dataspace and visualisation in Figure B-1, Channel 2 in Figure B-7, and the middle channels in Figure B-9.

Alerting, in contrast, requires no information transfer to the user until an alerting condition is detected. The information channel requirements are internal to the computer. In effect, the alerting daemons act as a very restrictive filter in the channel dataspace ⇒ display. They require very little actual display, but when display is required, it needs to be fast. The requirement is for high availability but low average bandwidth and low static entropy of the alerting display when it uses any display space at all. Alerting normally does not influence the channel display τo visualisation.

Monitoring or controlling concerns variation in the very low entropy element of the real world that is being monitored, in contrast to the high entropy world available for exploration. In Figure B-1, it is an end-to-end channel requirement, in Figure B-7 it is Channel 1, and in Figure B-9 it is the top-level channel. The display need not be more complex than is required to show the dynamic attributes being monitored in a minimal context. In other words the display can be of low entropy. The channel capacity required is low, but it needs to be continuously available. This requirement is almost the opposite of the requirement for alerting.

Searching is the most difficult perceptual mode for which to characterise the informational requirements. While search is in progress, the activity is very similar to Exploring. But the Search stops when the desired structure has been located. Search, then, would seem to require a relatively high-entropy display, that could change to suit the basic monitoring-controlling mode when the Search has succeeded. Search also needs high availability of the channel and a high bandwidth to allow for rapid changes of the region of the dataspace being Searched. Search, like Exploring, concerns the part of the channel from dataspace to visualisation.

These requirements are summarized in Table 2-2, copied here as Table B-3.

Table B-3: Informational Implications of Modes or Perception

	Availability	Instantaneous Entropy (Display Complexity)	Average Bandwidth
Monitoring/Controlling	High	Low	Low
Searching	High	High	High
Exploring	Low	High	Low
Alerting	High	Very Low	Very Low

B.3.3 Varieties, Causes, and Mitigation of Uncertainty

Above, we considered "uncertainty" as an abstract quantity that enters into information-theoretic constructs. According to Shannon [10], the information gained by a receiver is the reduction, upon receipt of a message, in the receiver's uncertainty about what the transmitter might have sent. The implication, contrary to what is often written about information theory, is that uncertainty and information are always somebody's uncertainty or information *about* something. Uncertainty and information do not exist in a vacuum, any more than do measures such as kilograms or centimetres.

In this section we consider different ways uncertainty might enter into the user's understanding of a network, and what to do about it. All the different kinds of uncertainty could, in principle, be used in information-theoretic analysis of networks, but we will not follow that conceptual path here. The following discussion is largely based on the results of a working group on uncertainty at the 2007 IST-059 Network of Experts Workshop (El Segundo, CA, USA).

B.3.3.1 Problem Definition

> *"It ain't what you don't know that gets you into trouble. It's what you know for sure that just ain't so."* (Mark Twain)

Another way of putting this aphorism is that misplaced certainty can be more damaging than recognition of uncertainty. The problem of representing data is as much to avoid inducing the user to believe falsely that something is true as it is to help the user to perceive what is really implied by the data. A good Framework should at least allow for the possibility that the reliability of data may need to be displayed in addition to its most likely value.

B.3.3.1.1 Uncertainty is in the User's Head

Like visualisation, uncertainty is in the head of a user. This uncertainty comes in two forms, which might be characterized as "I am uncertain of my understanding about X given these data" and "I understand that the data about X is imprecise". One refers to the person's appreciation that something in the world could be better understood, whereas the other refers to inadequacy in the data, an inadequacy that the person may or may not appreciate.

The world is what it is, but it can never be exactly known to anybody. Every observation has some imprecision. All we can ever have to work with is a "best bet" as to what is "out there", and this applies equally to the data supplied to the computer from external sources. Sometimes that "best bet" is treated as definite knowledge, but what that really means is that it is good enough for the purposes of the user. Besides the inherent uncertainty of any observation, there is a second kind of uncertainty in the data, which is the user's perception that more needs to be discovered about something. The user may perceive that some information is missing, or insecure. Perhaps different sources give different information about something on which they might reasonably be expected to agree. For example, one agent on the ground may claim that a bridge is intact, whereas another reports that it has been destroyed. The person receiving that information might then be uncertain about the state of the bridge, even though each source defines it precisely.

The other type of uncertainty can exist even though the data may be precise and complete. The data may be sufficient in principle to permit certain conclusions to be drawn, but the user may be uncertain as to how to draw those conclusions, perhaps because of inadequate training, perhaps because the data are too profuse, perhaps because the user failed to notice some key aspect of the data.

Suppose that the network data is accurate and complete, and is presented completely. A user might nevertheless be uncertain about its implications. As a trivial example, suppose that the data show that a particular structure forms a triangle having sides of length 3 m, 4 m, and 5 m. One user might realize immediately that this triangle has a right-angle between the sides of length 3 and 4, whereas another might be uncertain as to the value of that angle. Perhaps the network is too big to be fully comprehended, perhaps the implications of the network structure, parameters, and traffic patterns are not easily deduced, perhaps the task requires seeking out particular forms of sub-net that are hard to identify. All these are uncertainties in the user's head about the network that can exist even though all the necessary data are accurately displayed.

To some extent, this kind of uncertainty can be mitigated by effective display design (Section 4), but no display can guarantee to eliminate uncertainty from the user's head. Since the reason for visualising a network is usually to determine the implications of the data, this kind of uncertainty may be the most important.

In the context of networks, missing, indefinite, or misleading data is irrelevant if it does not affect the user's visualisation and understanding of the relation of the network to the real task. The problem, then, is that uncertainty affects the user's ability to make correct decisions in the task world, and to feel appropriately confident in those decisions. The real uncertainty is about the decisions. Uncertainty about the data and its implications is irrelevant if it does not affect uncertainty about a decision.

B.3.3.1.2 Uncertainties may be Inherent in the Data

Any network representing real-world data is inherently incomplete, if only in that there is an indefinite number of ways the nodes and links might interact with their real world surroundings, and most of those ways will not be represented in the abstraction of the network in the computer dataspace. Some of them may be related to the semantic embedding fields of the network, but the majority usually are not. The effects of abstraction may introduce uncertainty in the user, at least as to the likelihood that inferences from the computerised abstraction will be accurately applicable in the real-world task.

Aside, however, from the inherent effects of abstraction implied by representing the world as a network, the data available to the computer may be incomplete or ill-defined in many ways. Nodes may be missing or mischaracterized, as may links. If, for example, nodes represent persons, node "John" and node "Bob" might refer to the same person, but be represented as distinct. Links represented in the computer may connect nodes whose real-world counterparts are unrelated. It may be known that an out-link exists for a node but not to which other node that link connects, as might be the case if it was known that John mailed a letter but not to whom the letter was addressed. These and many other deficiencies of the network representation of the real world can lead the user to be uncertain about the behaviour of the network, about causal relationships linking events in one part of the network to consequences in another, or about prediction of future developments in the structure of, or in traffic over, the network.

B.3.3.1.3 Trustworthiness and Uncertainty

Above, it was noted that the user may be uncertain about various properties of a network even when it is a completely displayed accurate representation of the real world relevant to the user's task. The problem is compounded when the accuracy of the representation or the reliability of the source is in question. If the network properties are displayed as certain when the data are actually indefinite, the user may make unwarranted inferences about the real world, and finding them to be unwarranted, may later mistrust accurate data, leading to unnecessary uncertainty. On the other hand, the representation of uncertainty in a display imposes a load on the user's attention, and may detract from the user's appreciation of important represented properties; though the user may have less reason for mistrust, her likelihood of making useful and confident inferences may be reduced.

The user needs to be able to trust that when an item is displayed as definite, it is likely to be so; on the other hand, an over-enthusiastic representation of degrees of data uncertainty can detract from the usability of the display. To display an attribute takes some display real-estate, and requires some attention from the user. To display the uncertainty of the attribute doubles those demands. If the certainty or uncertainty does not matter, then it probably should not be displayed. It is in following chains of causality that uncertainties are likely to matter, because they propagate. If the task involves creating and following such chains, it would probably be useful for displays to show a parallel representation of how the uncertainty is likely to propagate – people are ordinarily quite prone to overestimate the certainty of their own inferences! Mark Twain's aphorism cited at the head of this section is most appropriate here.

B.3.3.1.4 When Uncertainty Matters

Uncertainty about some state of the world does not matter if, so long as the reality is within the range of uncertainty, the user's actions will be the same no matter what the correct data may be. It follows, then that in most circumstances uncertainty should be represented only if different data values within the range of uncertainty will lead the user to different decisions. If it matters whether a link from Joe Doakes connects to Stan Smith or to Stan Jones, and the data are ambiguous in that respect, then the dual possibility should be shown and highlighted. If it does not matter, then it may be immaterial whether either, both, or neither is shown. The uncertainty, as such, should not be shown in a way that either clutters the display or that might distract the user's attention from the material that could influence decisions.

The problem with this comment is the implication that the representational system knows the user's requirements. That is a difficult problem to resolve, except in an interactive display and possibly a mediated display (Section 4). Furthermore, it is easy to imagine a situation in which the user does not realize that some data are uncertain in a way that would influence her decisions, and in which the user has no reason to request the system to display the actual uncertainty for the critical data. Nevertheless, it remains true that uncertainty matters only when variation of the data within the range of uncertainty would affect the user's understanding of the situation in a way that would influence the user's decisions.

B.3.3.2 Reasons for Uncertainty

One cannot be uncertain about something of which one has no knowledge. It is, however, possible to be wrong, if something exists of which one has no knowledge. The "Uncertainty" Working Group at the 2007

N/X Workshop chose to include being wrong about the world among their types of uncertainty, and it is reasonable that ways of being wrong should be included in the Framework, even though in principle they cannot be represented in a display.

In the context of networks, one can be wrong from lack of data or from having incorrect data that one believes to be true. In a social network, an example of lack of data might be ignorance of the fact that Joe regularly communicates with Stan, which would mean that the network failed to include a link between two nodes that should have been linked. Such a link could well be crucial in analyzing terrorist network interactions. Incorrect data, on the other hand, might arise from assuming that Joe's communications to Stan were to Stan Smith, whereas they had actually been to Stan Jones. In this case, the link would exist, but would be misplaced in the network.

Being wrong is not the same as being uncertain, and wrongness is not the same as imprecision. If the data are wrong, no method of display can represent their wrongness. What might, however, sometimes be implicit in a representation is inconsistency. If some data are wrong, the implications of one part of the dataspace may conflict with the implications of another part, suggesting that one or other is in error. Bishop Berkeley attempted to demonstrate the reality of the world by kicking a rock he could see. Had his kick met thin air, his visual and tactile data from the world would have been inconsistent, and he would have had cause to wonder whether the visual rock had been an illusion or whether his kick had been misdirected. Inconsistency in the implications of data can be as much a source of uncertainty as is inconsistency of the data itself.

What attributes of uncertainty might be usefully displayed? The N-X working group at the El Segundo meeting listed Reliability, Confidence, Accuracy, Precision and Consistency. These attributes refer to different components in the train that leads to confidence in a decision.

- Reliability refers to the source of data and the route between that source and the data as displayed. It has a historical background, since a source cannot be known to be reliable or unreliable from one report. Only after several reports have been received and their data checked against other data from the same or different sources can the reliability be assessed.
- Confidence may refer to the confidence of the source in the data or to the confidence of the user in the data or in the implications of the data.
- Accuracy might refer to the correctness of the data as compared to the real-world truth, but since this can never be ascertained, it is not a very useful construct. It is possible, however, to assess the likely range of deviation of a particular datum from what might be the result of other measures of the same thing, and it is not unusual for a measure to be given as $x \pm y$.
- Precision refers to the likelihood that successive measures of the same thing result in similar data. Both Accuracy and Precision are more readily considered in connection with an attribute that has a scalar or vector value than in connection with the structural attributes of a network.
- Consistency refers to the repeatability of a datum based on different observations of the same thing. In a network context, this might include such things as whether A and B are connected by a link on Tuesday if they were on Monday. In this sense, Consistency has a wider range of application than does Precision.

Other than Confidence, all these possible attributes of uncertainty relate to the provision of the data for presentation to the user.

B.3.3.3 Consequences of Uncertainty

Mark Twain's aphorism at the head of this section warns of the danger of misplaced confidence, but lack of confidence in the data or in the implications of the data has its own danger. If the user is uncertain about the implications of data, decisions might be delayed when inaction is more dangerous than any of the uncertain possible choices of action. The user presumably will make those decisions after an appropriate risk-benefit analysis, whether that is done explicitly or intuitively. In battle and politics, this is the usual state of affairs, and there is little that can be done with display techniques to change that fact. Skilled users can make effective decisions with less overt information than is needed by novices or trainees.

If data that is imprecise or ambiguous is displayed in the same way as well attested data, even an experienced user may make a wrong decision. On the other hand, if the uncertainty of data is displayed too obtrusively, the user might be tempted to postpone a decision, an outcome that could be as dangerous as making the wrong decision.

A different kind of consequence of uncertainty is stress on the user. The less certain the user is about a decision that must be made, the more stress the need to make the decision is likely to cause. However, uncertainty about the decision is not necessarily increased by representing dubious data as if they were certain. Anomalous data can reduce the user's certainty about the interpretation of the displayed situation, whereas to represent correctly that the data are not reliable may make it easier for the user to discount the anomalous possibility inherent in the range of potentially correct values for the data. If Bob has been communicating frequently with Joe, it may be very important that a particular message was sent to Stan, but if the supporting information is unreliable and the message might well have been another in the series sent to Joe, it would be unfortunate if it were to be displayed as having been certainly sent to Stan.

As noted earlier, for the purposes of display it is important to know whether the user's decisions or situation awareness would be changed if the displayed element varied over its range of uncertainty. It is not possible for any automated display to have that information, so if any advantage is to be taken of this observation, it can only be in the context of interactive or perhaps mediated displays. When the user interacts with a display, it should be possible to request a display of the maximum likelihood situation, as well as to investigate extreme possibilities that are consistent with the information stored in the computer. However, the user is likely to take advantage of this only if the possibilities are made evident by some indication of the fact that certain aspects of the displayed data are uncertain, or if there are critical elements in the display for which the user knows that uncertainty might affect a consequent decision, and for which the user is able to query their credibility.

B.4 DISPLAY, USERS, AND INTERACTION

This section is not specific to displays of networks, but applies to all kinds of information display for single or multiple users. Nevertheless, the considerations introduced here should be an element of any complete Framework for network display. The language used in this section is that of visual display, but it should always be kept in mind that displays for visualisation may not themselves be visual.

Not all displays are intended for manipulation by the final recipient of the displayed information. A single operator who is also the end-user can interact freely with the display to influence what is displayed and how it is displayed, to the extent permitted by the hardware and software. This is the situation implied by the VisTG Reference Model (Figure B-7a) As soon as another person is involved, whether it be a single end-user who interacts with a display operator, or a large group such as an audience at a briefing or the readership of a book, the display cannot so readily be manipulated to suit all those who might need the displayed information.

If there are multiple users, especially if they are not viewing simultaneously, viewing is likely to be purely passive, which means that the "Controlling" mode of perception is not used. "Monitoring" is possible (and by implication, "Searching" also), but the "Exploring" mode becomes dominant. In Exploring mode, the user is learning about the dataspace so that the information learned will be available for later possible Controlling or Monitoring. In the case of networks, this means that the multiple users are likely to be examining the structure of the network rather than any ongoing activity in it. Even when they are not Exploring, but are Monitoring a common display, most, if not all, of the several users are unable to influence the content or nature of the display, which implies that the Alerting mode will seldom be appropriate.

We identify four different ways in which displays can be used: Interactive, Coordinated, Mediated, and Passive. In the first three modes, the display characteristics are altered in real time while the display is being used.

- In Interactive mode, a single end-user manipulates the display and the database in real time. This is the canonical situation reflected in the VisTG Reference Model of Figure B-7a.
- In Coordinated mode, more than one end-user observes the display, and more than one user has responsibility for altering the display. The coordination among the users is an issue, and only one of them can be controlling any one aspect of the display at a given moment.
- In Mediated mode, one person, whom we may call an operator or a presenter, interacts with a display on behalf of the end-user. A lone user such as a commander might be able to ask the operator to change the display in this way or that; in a briefing situation any of the viewers may be able to ask the person doing the briefing about aspects of the displays. In either case, there is interaction between the mediator and the user(s), as well as between the mediator and the display.
- In Passive mode, the user observes the display without influencing it in real time. An unlimited number of users can observe any particular display in passive mode. The display itself may change under the influence of an operator, but the users have no influence on the operator.

Table B-4 suggests which perceptual modes are most likely to be used under different circumstances. Some modes are not applicable under some circumstances. A single user cannot be working coordinated, as coordination implies that more than one user is actively observing and influencing the display; multiple users viewing simultaneously cannot all be interactively controlling the display or its content; and if multiple viewers look at the display at different times and places, the display is very probably static, which implies passive viewing. Most often, the only effective perceptual mode for passive viewing is Explore. The user looks to see what can be discovered about what is displayed, and expects it to remain valid for some time thereafter.

Table B-4: Perceptual Modes Most Likely to be Used in Different Circumstances

	Interactive	**Coordinated**	**Mediated**	**Passive**
Single End-User	All Modes	N/A	Explore, Search	Explore
Multiple Users Viewing Simultaneously	N/A	Monitor, Explore, Search, Alert	Explore	Explore
Multiple Users Viewing Separately	N/A	Monitor, Explore, Search, Alert	N/A	Explore

B.4.1 Single User

The VisTG Reference Model includes a complete outer loop, around which the user understands and influences the contents of the dataspace. The implication is that the user may interact with the display, or at least with the data being displayed. The model also has an intermediate loop around "Visualising" and "Engines". When the VisTG Reference Model is seen in its MVC abstraction, this is the loop that implements the Controller and the View, both being the responsibility of Engines, while the dataspace contains the MVC Model. In this intermediate loop the user visualises the implications of the display and influences the choice of data and the method of display. Again the implication is that the user may interact with the navigational and display Engines that shape what is actually displayed.

One problem with this approach is that not all displays are presented interactively. Displays in a book, a PowerPoint briefing or (usually) a Web page, are created by someone removed from the end-user in time and place. The whole display is available for the end-user's perusal, but the user cannot affect what is displayed. Almost all displays presented as examples of technique are of this kind. Such displays must be viewed passively, and passively viewed displays impose different requirements on the display syntax than do interactive displays.

When a user is interacting with a display, the relationships among display elements are naturally brought to the user's attention in their turn, as the user performs various manipulations in addressing the task. In contrast, in a passively viewed display, those relations are all presented simultaneously to a viewer who cannot manipulate the display to clarify ambiguous or non-obvious relations. The syntax of the display presentation must guide the user, in the same way that the syntax of written language guides the reader to perceive the relationships among the words, whereas in interactive conversation each partner can query ambiguities, and a less formal syntax is normal.

Mediated viewing comes in two flavours. In the first, an end-user does not manipulate the display directly, but asks an operator to generate the desired changes in it. Often this is done because the operator has greater skill in manipulating the display than does the end-user, or because the end-user has other responsibilities that preclude spending time to interact with the display. A mediated single user has much poorer control over the display than does an interactive single user, simply because the operator has to understand the end-user's intent in order to change the display appropriately. Because of this reduction in the user's control, a mediated display ordinarily will require somewhat more formal syntax than will a fully interactive display, though probably not as formal as is required for a static passive display.

The second form of mediated usage occurs in briefing. One person generates the displays in order to convey information to one or more others. The distinction between this and the first form of mediation is that in briefing, the controller of the display is the primary determiner of what is to be displayed. The person or group being briefed may observe passively for the most part, but sometimes they may be able to ask the briefer to bring out certain aspects of the information. In that case the situation is a mediated interaction. Mediated display thus shades without a clear boundary between being a substitute for single-person interaction, through the on-line briefing situation, to the construction of displays for passive viewing. The core situations are distinct, but the boundaries among them are fuzzy.

Just as the character of a good display will depend on the data structure, the task, the background knowledge and ability of the viewer, the static or dynamic nature of the display, and the perceptual mode (Controlling/ Monitoring, Searching, Exploring, or Alerting), so it will also depend on whether the manner of using the display is interactive, mediated, or passive.

B.4.2 Multiple Users

Most of the discussions in this report and elsewhere concerns displays for single users, despite that many displays are intended to provide information to many different users, although perhaps only one at a time. Pictures in books, on scrolls, or on clay tablets are the oldest and best understood multiuser displays, but multiuser displays are also prominent in computer-driven representations. They are likely to be passively viewed, though interaction is possible, especially in a mediated form. For example, in a briefing to a small audience, audience members may be able to ask the presenter to expand a portion of the display, or to show a different view.

A coordinated interaction with a display may occur, for example, in planning sessions, when several officers each may have the right to influence a display that reflects many aspects of a plan, and is seen by all. The term “coordinated” refers to the need to ensure that no two of the users influence the same aspect of the display at the same time. In coordinated interaction, the several users may communicate only through the mutually viewed display, or they may have other channels of communication, such as speech, to aid both in the coordination and in the interpretation of what is displayed.

The primary difference between a display intended for a single user and one intended for multiple users is that a single-user display can be tailored to the user’s expertise and background knowledge, whereas a multiuser display must take into account the different possible backgrounds and abilities of the target audience. Furthermore, unless all the users are simultaneously present, the display cannot easily be manipulated for them by a mediator. Hence, in most cases the only plausible mode is passive viewing in Explore mode.

The dichotomy between single and multiple users does not consider the case of serial single-viewer operation, such as when one air-traffic controller takes over from another. In that case, although the basic display design must take into account the different backgrounds of the several users, it is probable the successive operators have similar training and ability level. Even if they do not, the fact that they singly use the display suggests that each has the possibility of tailoring it to suit, rather than being required to accept design decisions imposed from elsewhere. Serial single-user systems need not be considered separately from pure single-user ones in Table B-4.

B.4.3 Passive versus Interactive Viewing

When viewing passively, the user is simply presented with a display. The display itself may change dynamically, but the user cannot influence it, even through the mediation of another person. Because most demonstration displays are in fact viewed passively, passive viewing is considered normal, and its consequences may not be obvious.

The difference between interactive displays and displays designed to be viewed passively is closely analogous to the difference between conversational language and text written for later reading. Written text must have a reasonably clear conventional syntax that identifies how the words relate within sentences, how the sentences cohere within paragraphs, and how the paragraphs combine to develop a theme. Conversational text is elliptic, single words or even facial expressions may substitute for what would be sentences in written text, and most importantly, the conversational partners can immediately query one another if they fail to understand the import of something the other said or did.

In the context of displays, displays intended for passive viewing must conform to some syntax generally understood by the target audience, and must include all the information necessary to make whatever point the display designer wants to get across. In contrast, interactive displays need only show enough to satisfy the momentary needs of the user, and need not be intelligible to anyone else – or to the same user at another time. If the interactive user does not understand the implication of the display and the interaction process is well designed, supporting information can be brought to play when it is needed.

That passively viewed displays must have a generally understood syntax is not to say they should be static. Many people have had the experience of listening to a lecture in which the lecturer covers a blackboard with equations, symbols, arrows, and boxes, and has found it easy to follow the flow of the lecture. A person with the same background, seeing the same blackboard later, might be totally unable to make sense of it. The temporal flow of the construction of the display is an important part of its syntax. In more contemporary technology, a good presenter using PowerPoint may build up a complex picture over a series of slides, adding and changing elements or using animation in an easily understood sequence. The audience will understand the complex result much better than if they were presented with it all in one static picture. Display syntax exists in space and time, together.

B.4.4 Coordinated versus Mediated Display

Coordinated and Mediated displays both are active, in the sense that the end-user has some influence on what is displayed, in real time. The means by which this influence is exercised differ. By “Coordinated” we imply that more than one person at the same time can directly manipulate the content of a display that is visible to all; by “Mediated” we mean that the end-user does not have direct control of the display, but exercises influence over it through the actions of an operator. Even in mediated presentation, multiple end-users can influence the same display, but the mediating operator performs the coordination.

Planning and team analysis are situations in which coordinated display is likely to be useful. In both cases, team members are likely to have different competences and roles, which implies that they will ordinarily want to manipulate different aspects of the display. One person may, for example, develop concepts for air attack routing, while another works out the logistical implications, both being displayed on the same screen. In social network analysis, one may highlight contact networks, another may seek family connections that could underlie the contact network, while yet another may examine resource availability on the assumption that the members of the network have some nefarious intent. All may affect both the Model (in MVC terms) and the Views available to all team members.

Teamwork using a common display can be supported either by coordinated or by mediated displays. The difference is that if the display use is mediated, the problem of deciding what display aspects to change and when to change them is given to a human operator. The human mediator may resolve conflicts, perhaps alone, or perhaps by pointing out to the users who have conflicting requirements that the conflict exists, thereby leading them to discover wherein their concepts differ, either about what is in the Model or about how it should be shown. In a Coordinated display without mediation, conflict must be resolved either by the use of side-channels such as voice for communication among the users, or by individual understanding of what is happening within the display. The former is more likely if the users are co-located than if they are looking at the same display on geographically separate screens.

Coordinated displays have syntactic requirements with a stringency that lies between the informality of interactive single-user displays and the structured nature of passively viewed displays. The coordinated display can be considered as a means of real-time communication among team members, allowing for immediate queries should one team member fail to understand the import of another’s display manipulations, thus reducing the need for a syntax that, in a passive context, would have disambiguated the manipulation. On the other hand, the team members often have different backgrounds and roles, which implies that display structures and manipulations must be at least a little more formal than they might be when a single user is interacting alone with the display. This difference in background and role becomes important when dealing with "The Common Operational Picture" (Section 4.5).

It is not necessary for all users of coordinated displays to be co-located. Indeed, they may not even see the same presentation. However, in coordinated display, what each user does to influence their local display may affect what the other users see on their own local displays, through the effects on the Model that they all have in their various Views. If there is no such cross-influence among the several displays, then the situation is not a coordinated display but a multitude of single-user interactive displays, possibly working on the same dataspace.

B.4.5 The Common Operational Picture

At the IST-043/RWS-006 Workshop on "Visualisation and the Common Operating Picture", Working Group 1 (WG-1 in the following) studied a structured approach to a purpose-driven Common Operational Picture. They argued [13] that the concept of a common picture was itself misleading, and that instead different team members needed to see how their mission objectives fitted into the common operational environment. In the Summary of their report, WG-1 said:

> *The notion of a COP has three components: "Common", which implies that there are at least two collaborating partners; "Operational", which implies that there is a real-time element involving action involving the partners; and "Picture", which implies that each partner has some kind of vision of the situation in which the action takes place. This report addresses the first two of these components. The "Picture" aspect involves for the most part issues that do not change between displays intended for one user and displays intended to facilitate the development of a vision common to two or more partners.*

Although, as discussed above, we might now query that last assertion, nevertheless different pictures would ordinarily be needed by the several team members in order that they all arrived at a common understanding. This holds as true for displays for network analysis and control as it does for the battle-planning or civil crisis displays that were the focus of the work of WG-1.

> *To create a "vision", a person integrates incoming data with memories and understandings already in the mind. The commonalities of background data can be enhanced by communications on widely different time scales, the immediately varying data being perhaps not very large, if the backgrounds are sufficiently similar (e.g. a blown bridge is easily described in a few bits of data, if the parties have a detailed reference map in common). This may seem self-evident, but it forms the basis of the proposed approach to the COP system. The underlying point is that when rapid cooperative action is required, very little data need be communicated between the partners, if they share (and know they share) an appropriate common background.*
>
> *In military systems, common backgrounds arise in several ways, not least through the medium of training in the doctrine of the services to which the partners belong. If they have training and "culture" in common, communication of a "Common" vision is much easier than if they belong to different services (in Joint operations) or to different nations (in Coalition operations).*
>
> *A primary issue with the notion of the "Common Operational Picture" is that for the picture to be "Common", the data on which cooperating parties base their picture must be up to date. This implies communication between the databanks on which the parties base their displays. The displays normally will not be in common, unless the parties are physically together, looking at (or listening to) the same display, but they should have sufficient commonality that each party understands the other's view of a situation well enough to be able to visualise how their respective roles in any action support one another.*

> *The nature of the operation in question is immaterial. Whether the domain be response to civil disaster, force-on-force battle, or delicate peace-keeping, the same questions of commonality of background and its effect on the amount of necessary data communication arise. Always there are the same two competing issues: the need for each partner to know that the other(s) know what they must if they are to cooperate effectively, versus the likelihood that the communication of data already known may obscure the reception of important novel data (the "clutter" problem).*
>
> *With these issues in mind, the working group developed a Venn diagram approach to the description of the information sharing among the partners. Much of this report discusses the implications of the elaborated Venn diagrams.*

For details on the development of the Venn diagram see the report of the Working Group [13]. Here we show an early and a late stage in its development, because they affect the design of displays, whether for networks or in the more general case.

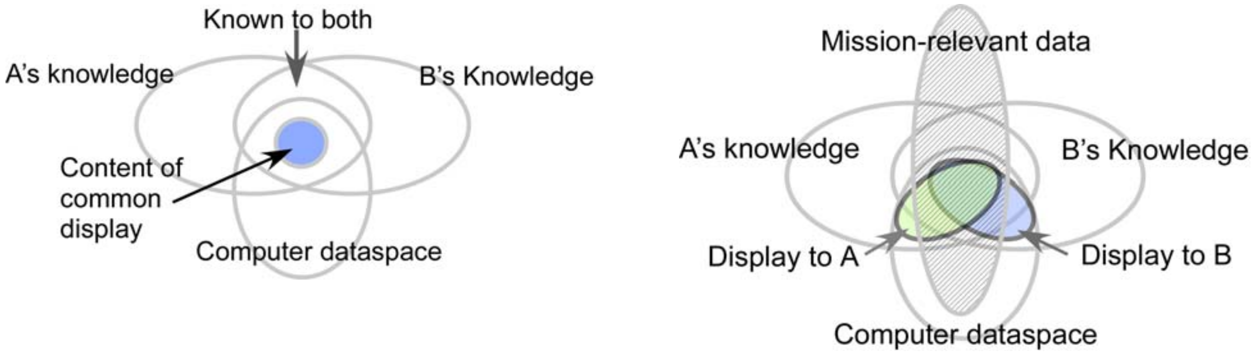


Figure B-10: The Basic Venn Diagram for a Coordinated Common Display for Two Users – (a, left) A and B both have independent but overlapping knowledge, which necessarily includes whatever is shown on the common display. The computer dataspace also contains information, only some of which is known to A or B; (b, right) Not all data relevant to the mission is available to A, or B, nor is it in the computer dataspace; A and B may be shown overlapping but distinct coordinated displays, in which not all the content is mission-relevant (after [13]).

The need to show that information is known to A or B indicates a significant problem with the Venn diagrams. They have no representation of whether A knows that B knows or does not know information that is shared (B does know it) or not shared (B does not know it). In other words, A and B may share knowledge of some information, which is therefore properly included in [the shared knowledge regions], but neither may realize that the other does know it. Alternatively, A may not realize that B does not know some critical item. The COP system should have some way whereby the partners can probe each other's understanding.

That the collaborating users should be able to probe each other's understanding is a critical point in designing coordinated displays, especially displays for planning. That they should be able to communicate reasons for making changes affecting what is displayed to the other(s) is independently important for most coordinated displays. The necessary communication facilities may be provided as part of the syntax of the displays (e.g. video chat panels in a corner of the screen) or may be independent of the displays, but they must exist. Ideally, the displays should incorporate some means whereby each user capable of influencing the displays seen by the others can indicate to the others the current goal of the operations.

B.4.6 The Display of Uncertainty

It is difficult to display both an attribute of a network or a network element and at the same time display the range of uncertainty about that attribute. The reason for the difficulty is that the same dimensions of display that are available to display the attribute are also those that are available for displaying its uncertainty. To use some dimensions for the display of uncertainty is to deny the use of those dimensions for the display of attribute values. If colour is used to show the uncertainty of some link, then colour cannot be used to show the traffic density on that link. If the display is varied over time to show the uncertainty of some value, then animation cannot be used to show the network dynamics. This issue is inherent: If a system has N values to be displayed, then to include the uncertainties associated with those N values requires 2N representation entities, plus the one-to-one linkages between each element and its uncertainty. Even if the display techniques permit 2N entities to be displayed, to display all the uncertainties would add considerably to the clutter and might well distract the user's attention form the important aspects of the display.

Having pointed out the inherent problem with displaying uncertainty, we can list some of the techniques that have been proposed or used. These include: Colours, Transparency, Blurring, Grey scaling, Glyphs or Symbols, Size, Thickness, Patterns, and temporal variation. All of those can also be used to represent attribute values. None, other than temporal variation, are suitable to indicate, say, a link for which one terminal is uncertain, or a node that might actually be the same individual as another node in the net.

Effective representation of uncertainty is an ongoing research problem, and it is not one that lies within the purview of IST-059/RTG-025. Here, we can do no more than note that the problem ranges from trivial to severe under different conditions. Although display designers may sometimes be able to find ways of representing uncertainty that can help a user who must base situation awareness on possibly uncertain data, they must always be aware that representing uncertainty is liable to distract the user from effective understanding of the implications of the data.

B.5 FRAMEWORK PROCESS AND WAY AHEAD

To use the Framework implies more than and yet less than knowing all the details of taxonomies and theories described above. It requires the user to imagine what a successful use of a display for the task at hand would mean, and then to make concrete whatever is known beforehand, thereby clarifying what needs to be found in the dataspace and displayed in a way that connects the new material to what the user already knew.

What does the user want to achieve? Simply to answer "Understand the network" is inadequate, even if the perceptual mode is Exploring. As discussed above, networks have many and varied aspects that might be of interest. The answer should specify whether the interest is in an overview for exploration purposes, a search for cliquish sub-nets or other localized aspects of the network, examination of possible dynamic modes of network behaviour, traffic analysis, or something quite different. In any case, the answer should specify the objective.

The variety of possible objectives is enormous, so the Framework process uses the taxonomies set out above in a way that should make it easier to specify the user's objective in terms that can be used in selecting effective display types or application software. Currently, the questions are set out in the form of a spreadsheet, as a help or aide-memoire for the user. As yet, the answers are not linked to anything else, such as suggestions for display characteristics. Those linkages are intended for further developments.

Further development of the framework will include mapping effective display types to patterns of responses in the worksheet. To do this will require two things. The first is that a variety of problem cases be walked through the spreadsheet, and that the issues that arise be used to evolve the spreadsheet into a more easily used form. The second is to compare the patterns of answers with display types that have been found to serve those problems well. This matching should be eased by a kind of dimensional reduction, using the display and data taxonomies, augmented to specialize in network representation by taking into account the network properties mentioned in Section 2 above.

The current worksheet, filled in for some different kinds of walk-through example problem, is shown in Chapter 6. These examples are for computer network protection, seeking evidence of whether a terrorist plot exists (set in the England of Elizabeth I), a contemporary terrorist social network, and a possible avian influenza epidemic. Both the context and the nature of the problem differ widely across the four examples, but it proved possible to give useful answers to the various questions that allow the problem to be characterised in a way that should lead to a useful selection of display techniques. This characterisation has not yet been done.

The Framework is intended eventually to be used with the survey, to find software or full applications that would suit the user's purpose. Chapter 5 describes the way this integration is conceived. To achieve it requires that the Survey be kept current with the ever-changing development of techniques and algorithms, and that the information be recorded in the database in a way compatible with the Framework taxonomies.

To keep such a database current is not easy, and nor is it easy for any one application's potentialities to be recorded by anyone not intimately familiar with the application. Nevertheless, to maintain some information in the database is more useful than to have none, and even if full integration of the Framework with the Survey were never achieved, yet each component has value in itself.

Implementation of all the potential implicit in the Framework is a daunting task. Figure B-11 suggests some of the disciplines that may be needed. Figure B-12 illustrates one view of how the Human Factors Engineering aspect (the central area of the VisTG Reference Model) might be developed.

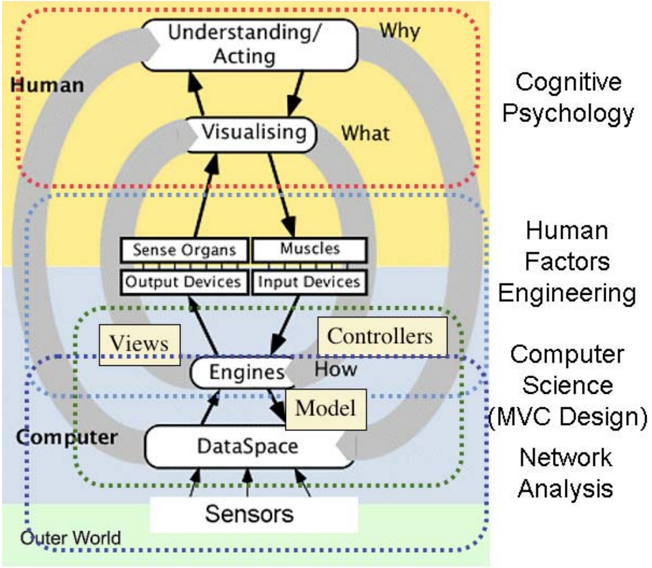


Figure B-11: Some of the Disciplines Involved in Developing the Potential of Elements of the Framework.

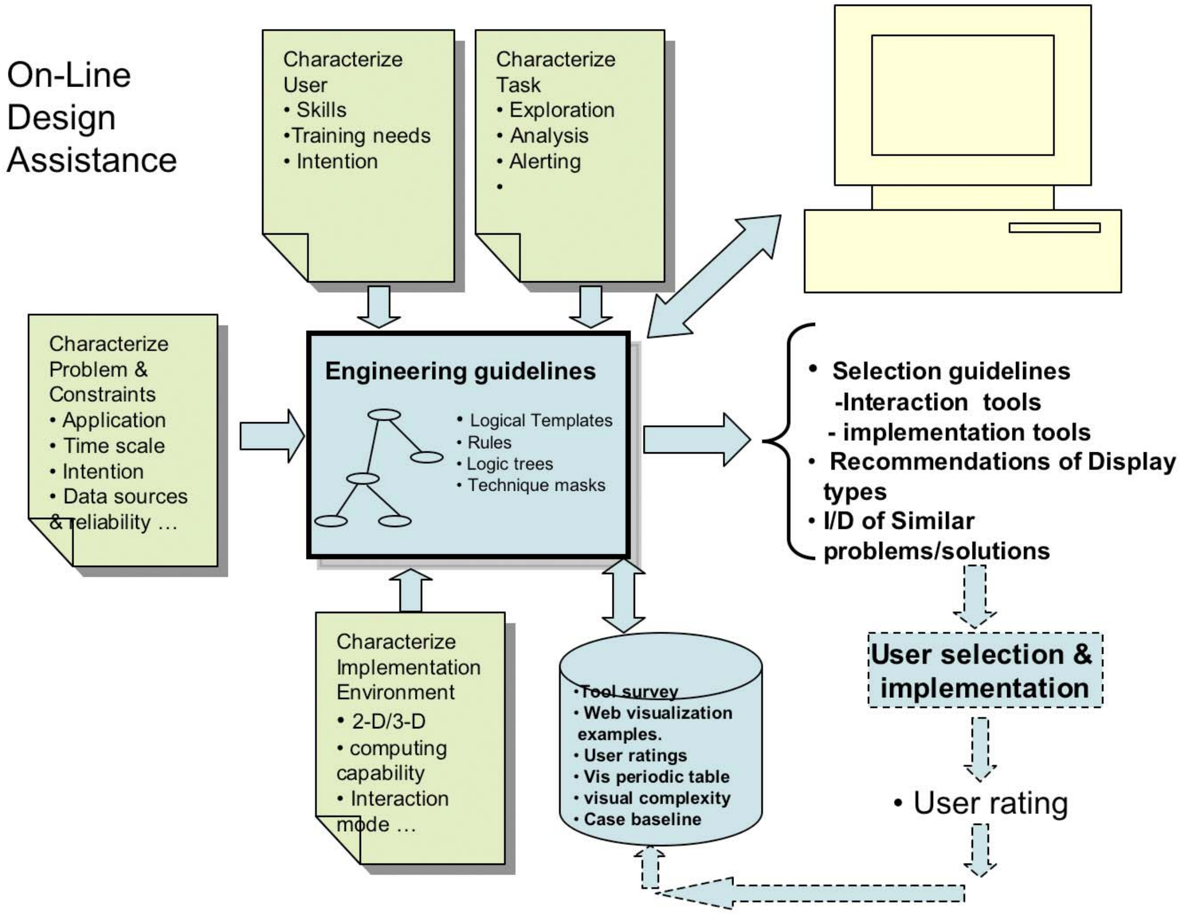


Figure B-12: A View of the Framework as a Human Factors Engineering Problem.

In summary, the way ahead has three paths in parallel. One is to enhance the worksheet, perhaps implementing it in software; the second is to map the pattern of worksheet answers to network properties and display types, and the third is to link them all in a software implementation to the Survey database so as to allow a user either to find suitable software or to determine that novel representations might be required. It is to be hoped that developments such as are implied by Figure B-12 will be pursued, as might similar proposals for the other disciplines mentioned in Figure B-11, but those developments are likely to be beyond the scope of any successor of IST-059.

B.6 REFERENCES

[1] Garner, W.R., “Uncertainty and Structure as Psychological Constructs”, John Wiley & Sons, New York, 1962.

[2] NATO IST-013/RTG-002, (2001), “Visualisation of Massive Military Datasets: Human Factors, Applications, and Technologies”, *Technical Report, RTO-TR-030*, NATO Research and Technology Organisation.

[3] NATO IST-043/RWS-006, (2005), "Visualisation and the Common Operational Picture", *Workshop Proceedings, RTO-MP-IST-043*, Toronto, Ontario, Canada: NATO Research and Technology Organisation, 14-17 September 2004.

[4] NATO IST-021/RTG-007, (2007), "Multimedia Visualisation of Massive Military Datasets", *Technical Report, RTO-TR-IST-021*, NATO Research and Technology Organisation.

[5] Caldarelli, G., "The Structure of Biological and Social Systems", *The Mathematics of Networks*, SIAM News, Volume 37, Number 3, April 2004.

[6] Erdös, P. and Rényi, A., "On Random Graphs. I.", *Publicationes Mathematicae*, 6: 290-297, (1959).

[7] Erdös, P. and Rényi, A., "The Evolution of Random Graphs", *Magyar Tud. Akad. Mat. Kutató Int. Közl*, 5: 17-61, (1960).

[8] Grassé, P.-P., "La Reconstruction du nid et les coordinations interindividuelles", *La théorie de la stigmergie, Insectes Sociaux*, 6: 41-84, (1959).

[9] Wright, W., "Visualisation for Sonar Tactic Decision Aids", Presented at *NATO Workshop Visualisation of Massive Military Multimedia Datasets*, Valcartier, Québec, Canada, 2000, NATO RTO-MP-050.

[10] Shannon, C.E. and Weaver, W., "The Mathematical Theory of Communication", Urbana, University of Illinois Press, 1949.

[11] Smestad, T., "Outline for Concepts and Principles of Effective Visualisation", *Annex A to RTO-TR-IST-021 – Final Report of IST-021/RTG-007 Multimedia Visualisation of Massive Military Datasets*, 2007.

[12] Garner, W.R. and Clement, D.E., "Goodness of Pattern and Pattern Uncertainty", *Journal of Verbal Learning & Verbal Behavior*, 1963, Vol. 2(5-6) 446-452.

[13] "Towards a Structured Approach for the Development of a Purpose Driven COP System", *Report of Working Group 1, in Proceedings of IST-043/RWS-006*, NATO RTO-MP-IST-043.

Annex C – SOCIAL NETWORK ANALYSIS
(or ANALYSIS OF CRISP POINT-TO-POINT NETWORKS)

M.R. Nixon

Most analytic studies of network properties have been concerned with crisp point-to-point networks abstracted from any possible embedding field. In other words, these properties are intrinsic to the networks concerned. This section lists a few of them, using social networks as the example type, since social networks have most of the attributes of any crisp point-to-point network.

C.1 ANALYTIC ABSTRACTIONS OF CRISP POINT-TO-POINT NETWORKS

A social network is a collection of nodes representing members or groups of an underlying population together with ties or links between these nodes denoting binary relationships. As such, a social network is a refinement of a semantic network in which nodes stand for socially significant entities such as:

- People
- Units of action
- Coalition partners
- Departments
- Resources
- Ideas or Skills
- Events
- Nation-states

While the binary relationships indicated by links answer socially significant questions about the nodes they tie:

- Who do you like or respect?
- Transfer of resources
- Authority lines
- Association or affiliation
- Alliance
- Substitution

This section presents some of the most useful properties of social networks as abstract semantic networks. We concentrate on mathematical definitions for semantic networks that support measurement of socio-cultural environments of interest. Our aim is to define measures on graphs (binary matrices) and networks (weighted matrices). What follows is a synopsis of presentations given on the subject by Professor Kathleen Carley of Carnegie-Mellon University and which she has graciously agreed to share with IST-059/RTG-025 [1]. Figures and tables have been provided by Professor Carley. Errors in this Annex are due to the author and not to Professor Carley.

C.1.1 Network Data

Data suitable for defining an underlying system of social relationships can be gathered from empirical trials, generated by simulation or simply stipulated. They take the following forms:

Table C-1: Systems of Relationships

Type	1-Mode	2-Mode
Dyad	Many actor-to-actor pairs	Pairs – actor-to-actor, location-to-location, actor-to-location
Ego	1 actor to other actors	1 actor to other actors and locations
Full	Actor-to-actor	Actor-to-location

The modes of a graph or network are its node types, e.g. workers, factories, tasks. Note that in a "Full" social network, we may have many separate cliques of actors which do not share members. The "Ego network" of one of its members is that member's clique in the full network. Location is merely one example of a mode other than actor in network data – task, time, or expense might be others. In general, a mode is a sub-class of the semantic network's universe of discourse other than actor which is also essential to measuring network behaviour of interest.

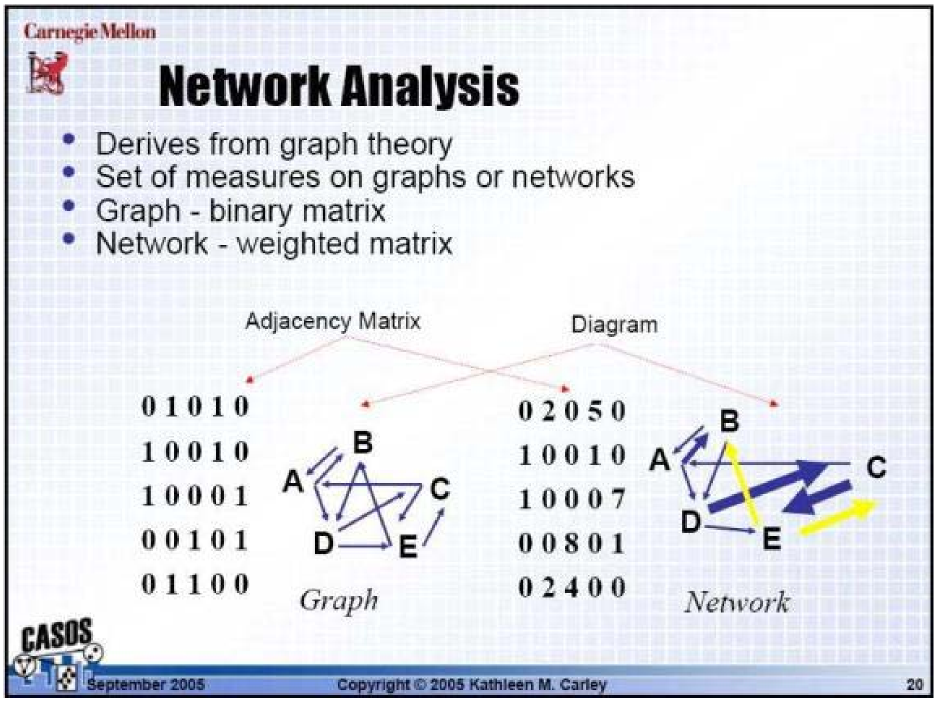
C.1.2 Graph and Network Representation

A graph or network is represented by a square matrix indexed in both dimensions by an enumeration of its nodes and containing entries for ties (links). As shown in the figure below, ties among nodes are indicated in the adjacency matrix for a graph or network as existing by virtue of non-zero entries – there being no distinction made between a missing tie and a known non-existent tie. Graphs or networks of undirected ties, e.g. "works with" or "is married to", are symmetric. Symmetry, we note, is defined by reflection about the diagonal of the adjacency matrix such that $R_{i,j} = R_{j,i}$.

For many key measures over symmetric graphs/networks, the diagonal of their adjacency matrices assume the value of 0, as shown in Figure C-1, for all node indices i so as to eliminate reflexive ties that might otherwise be indicated, e.g. whether actors "work with" themselves. This figure illustrates:

1) That the basic mathematical entity is the graph as represented by a binary (Boolean) adjacency matrix indicating the presence or absence of a link; and
2) A network is a graph in which nodes are linked by capacity, exchange or some other (real, not Boolean valued) measure of value.

For that matter, networks can have dynamic as well as non-scalar link values. Bayesian belief networks (BBNs), for example, have links directed at nodes according to conditional dependency. Each parent node in a BBN exchanges transient probability distributions (tables or functions) with its child nodes over their links to one another during the process of inferring new posterior probability distributions.



Carnegie Mellon

Network Analysis

- Derives from graph theory
- Set of measures on graphs or networks
- Graph - binary matrix
- Network - weighted matrix

Adjacency Matrix

Diagram

0 1 0 1 0
1 0 0 1 0
1 0 0 0 1
0 0 1 0 1
0 1 1 0 0

Graph

0 2 0 5 0
1 0 0 1 0
1 0 0 0 7
0 0 8 0 1
0 2 4 0 0

Network

September 2005

Copyright © 2005 Kathleen M. Carley

20

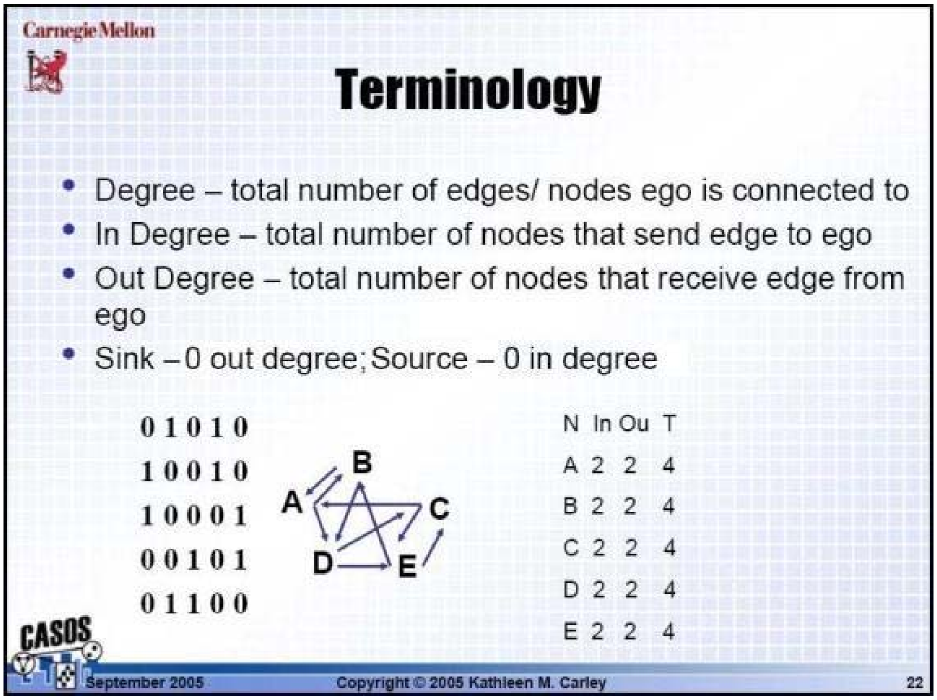
Figure C-1: The Adjacency Matrix and Visual Representations of a Graph and a Network (in this section, the embedding fields of networks are ignored, and a network is taken to be a graph with weighted edges).

The strength of a tie/link in a network, indicated by its weight, is abstract enough to represent quantities as diverse as how frequently actors represented by its nodes interact with one another or, again, the distance or cost of transportation between two locations. However, as noted in Section 2.1.3.1, “weight” can have a variety of implications, at least for a traffic-bearing link. It could mean any of “capacity”, “utilization”, or “availability”, at least.

With sequences of ties/links, we distinguish a walk (an unrestricted sequence of ties between adjacent nodes) from a path (a walk in which no node is visited more than once) and both from a trail (a walk in which no tie is repeated). Paths starting from one node and ending in another are used to define the distance between the two nodes as the number of ties in the shortest path (geodesic) joining them.

Other key graph-theoretic concepts include the distinction between directed (a commands b) or undirected (a works with b) ties/links. Directed ties are those indicated in the adjacency matrix by entries $R_{i,j} \neq 0$ and $R_{j,i} = 0$ thus failing the condition for being a symmetric graph/network presented earlier. Here we mention the concept of being a transitive graph/network which satisfies the condition that if $R_{i,j} = R_{j,k} \neq 0$ then $R_{i,k} \neq 0$. Transitivity is a critical property of any packet-switching router network for the correct forwarding of multilevel secure data, i.e. routers need to have authentication ties/links to those serving a packet’s destination even if the packet is to be routed through intermediates before arriving at its destination in several hops.

The degree of a node in a graph or network enters into the definition of many of its other attributes and measures. Figure C-2 provides a tabulation of the degrees (in-, out- and total) for the nodes in our simple graph from the previous figure. It reveals that no node in the graph is a source or a sink.



Carnegie Mellon

Terminology

- Degree – total number of edges/ nodes ego is connected to
- In Degree – total number of nodes that send edge to ego
- Out Degree – total number of nodes that receive edge from ego
- Sink – 0 out degree; Source – 0 in degree

```
0 1 0 1 0
1 0 0 1 0
1 0 0 0 1
0 0 1 0 1
0 1 1 0 0
```

N	In	Ou	T
A	2	2	4
B	2	2	4
C	2	2	4
D	2	2	4
E	2	2	4

CASOS

September 2005 Copyright © 2005 Kathleen M. Carley 22

Figure C-2: The Degree of a Node in a Graph or Network is the Total of its In- and Out- Links.

C.1.3 Overview of Graph and Network Measures

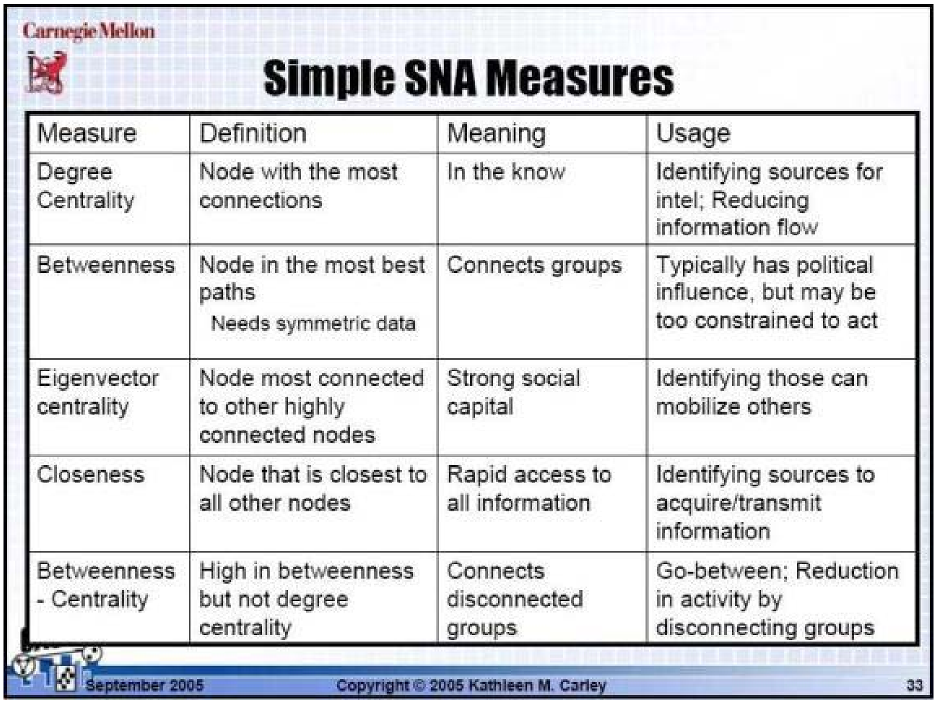
Measures at the level of a graph or network include its size, i.e. number of nodes in contains and its density, i.e. the ratio of the number of actual ties to the number of possible ties between nodes involved. At the dyadic level the link or tie, we measure frequency, i.e. for a given link the ratio of the number of distinct paths in the network passing through that link to the total number of distinct paths. At the node level, we measure centrality, i.e. relative degree (in- or out-), of a node so as to discern, e.g. key actors such as heads of hierarchies. Coming full circle, at the level of an entire graph/network, centralization then indicates the extent to which the graph/network is focused on a single node (or set of structurally equivalent nodes).

C.1.4 Graph/Network Level Measures

We measure the extent of a property like symmetry or transitivity in a graph/network as the ratio of the numbers of pairs (triples) of nodes that satisfy the symmetry (transitivity) condition to the total number of pairs (triples) from the graph/network.

C.1.5 Node Level Measures

Figure C-3 below summarizes the main node level measures used in Social Network Analysis.



Carnegie Mellon

Simple SNA Measures

Measure	Definition	Meaning	Usage
Degree Centrality	Node with the most connections	In the know	Identifying sources for intel; Reducing information flow
Betweenness	Node in the most best paths Needs symmetric data	Connects groups	Typically has political influence, but may be too constrained to act
Eigenvector centrality	Node most connected to other highly connected nodes	Strong social capital	Identifying those can mobilize others
Closeness	Node that is closest to all other nodes	Rapid access to all information	Identifying sources to acquire/transmit information
Betweenness - Centrality	High in betweenness but not degree centrality	Connects disconnected groups	Go-between; Reduction in activity by disconnecting groups

September 2005 Copyright © 2005 Kathleen M. Carley 33

Figure C-3: Node Level Measures.

We define the varieties of centrality for measuring nodes according to Carley's presentation.

First, we define **Degree Centrality** (Extraversion)

$$C_D(v_k) = \sum_i \frac{deg(v_k)}{deg(v_i)}$$

where $deg(v_k)$ is node k's degree. Degree centrality is one of the simplest measures of a node's significance to understand and certainly one with intuitive visual content. A node can be distinguished by grey-tone or transition colour in displaying its degree centrality.

Next, we define **Betweenness**. Vertices that occur on many shortest paths between other vertices have higher betweenness than those that do not. For a graph $G = (V,E)$ with n vertices, the betweenness $C_B(v_k)$ for vertex v_k is:

Betweenness Centrality (Influence)

$$C_B(v_k) = \sum_{s \neq v_k \neq t \in V} \frac{\sigma_{st}(v_k)}{\sigma_{st}}$$

where σ_{st} is the number of shortest geodesic paths from s to t, and $\sigma_{st}(v_k)$ the number of shortest geodesic paths from s to t that pass through a vertex v_k. This may be normalised by dividing through by the number of pairs of vertices not including v_k, which is $(n - 1)(n - 2)$. Like degree centrality, betweenness centrality also has straightforward intuitive visual content as an indicator of how greatly modifications (e.g. isolations by tie severing) to a node will affect overall graph/network behavior and in a display can also be distinguished visually by grey-tone or transition colour. Importantly, one may also visualise simultaneously (or by selection) the betweenness of a node by colouring distinctively all and only those geodesic paths passing through it.

Closeness is another centrality measure of a vertex within a graph. Vertices that are 'shallow' to other vertices (that is, those that tend to have short geodesic distances to other vertices with in the graph) have higher closeness. Closeness is preferred in network analysis to mean shortest-path length, as it gives higher values to more central vertices, and so is usually positively associated with other measures such as degree. The closeness $C_C(v_k)$ for a vertex v_k is the reciprocal of the sum of geodesic distances to all other vertices in the graph:

Closeness Centrality (Access)

$$C_C(v_k) = \frac{1}{\sum_i d_G(i,k)}$$

Eigenvector centrality is a fourth measure of the importance of a node in a network. It assigns relative scores to all nodes in the network based on the principle that connections to nodes having a high score contribute more to the score of the node in question. Using the adjacency matrix to find eigenvector centrality, we let x_i denote the score of the ith node, v_i. Let $A_{i,j}$ be the adjacency matrix of the network. Hence $A_{i,j} = 1$ if the ith node is connected to the jth node, and $A_{i,j} = 0$ otherwise. For the ith node, the Eigenvector centrality score is proportional to the sum of the scores of all nodes which are connected to it:

Eigenvector Centrality (Status)

$$C_E(v_i) = x_i = \frac{1}{\lambda} \sum_{j \in M(i)} x_j$$

(where $M(i)$ is the set of nodes that are connected to the ith node, N is the total number of nodes and λ is a constant) or equivalently using the adjacency matrix,

$$x_i = \frac{1}{\lambda} \sum_{j=1}^{N} A_{i,j} x_j$$

in vector notation this can be rewritten as:

$$\overrightarrow{x} = \frac{1}{\lambda} A \overrightarrow{x} \text{ or, } A\overrightarrow{x} = \lambda \overrightarrow{x}$$

which is the eigenvector equation. Hence the ith component of the eigenvector corresponding to the eigenvalue λ gives the centrality score of the ith node in the network.

Google's PageRank is a variant of the Eigenvector centrality measure.

Dependence centrality is a fifth measure of the importance of a *node* in a *network* described in [2]. It is a node-level measure based on a network-level measure of *efficiency* $E(G)$. Network efficiency is a measure quantifying how efficiently the nodes of the network exchange information. To define efficiency of G, we first calculate the shortest path lengths d_{ij} between two arbitrary nodes i and j. We now suppose that every vertex sends information along the network through its edges. The efficiency ϵ_{ij} in the communication between vertex *i* and *j* is inversely proportional to the shortest distance: $\epsilon_{ij} = 1/d_{ij}$, $\forall i, j$. When there is no path in the graph between *i* and *j*, we get $d_{ij} = +\infty$ and consistently $\epsilon_{ij} = 0$. *N* is known as the size of the network or the numbers of nodes in the graph. Consequently the average *Network Efficiency* of the graph of *G* can be defined as:

$$\begin{aligned} E(G) &= \frac{\sum_{i \neq j \in G} \epsilon_{ij}}{N(N-1)} \\ &= \frac{1}{N(N-1)} \sum_{i \neq j \in G} \frac{1}{d_{ij}} \end{aligned}$$

The above formula gives a value of *E* as a fraction of unity. From this network-level measure of efficiency, Memon and Larsen derive *Dependence Centrality* which measures how much a node is dependent on any other node in the network:

Dependence Centrality (Hierarchy)

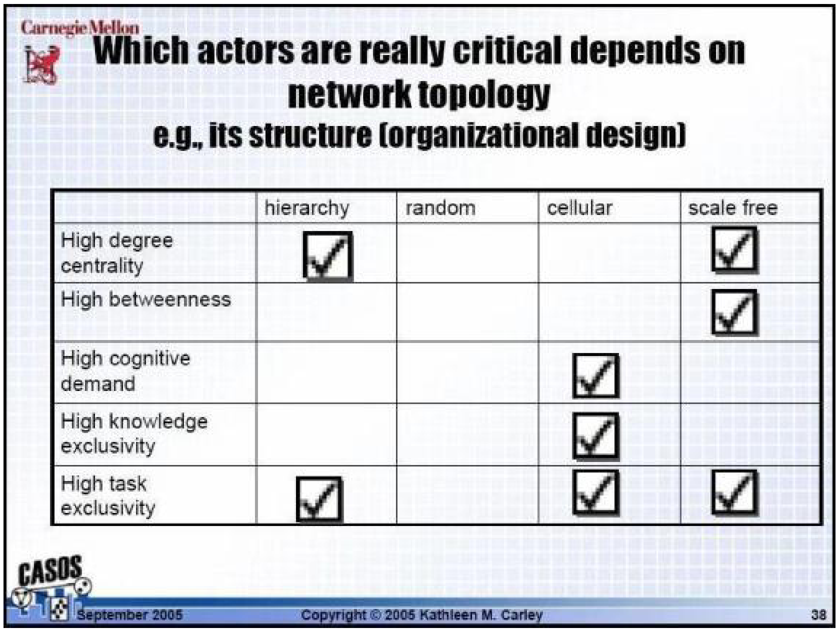
$$DC_{mn} = \sum_{m \neq p, p \in G} \frac{d_{mn}}{N_p} + \Omega$$

where *m* is the root node which depends on *n* by DC_{mn} centrality, N_p is the actual number of geodesic paths leading from *m* to *p* through *n*, and d_{mn} is the geodesic distance from *m* to *n*.

Note that Ω is taken to be 1 if graph is connected and 0 in case it is disconnected. In [2], the authors take Ω to be 1, because it is assumed that the graph is connected. The first part of the formula tells us how many times *m* uses *n* to communicate with other nodes *p* of the network. In other words, *p* is any node of the network to which *m* is connected through *n* (the connection represents the shortest path from node m to *p* with *n* in between). N_p represents the number of alternatives available to *m* to communicate to *p* and d_{mn} is the multiplicative inverse of geodesic distance 1/*d*.

C.1.6 Graph/Network Measures and Topology

The topology or structural organization of a graph/network influences how meaningful measures such as closeness, betweenness and the centralities are in describing it. In a complete (fully connected) network, all nodes have equal in, out and total degree. Indeed, they also have equal average betweenness, closeness, etc., in relation to all other nodes. So, none of the measures introduced so far will distinguish nodes in a complete graph/network. Non-trivial (incomplete) graphs/networks have many missing ties/links and assume, thereby, different topologies according to those remaining. Figure C-4 shows a number of different topologies and various graph/network measures suitable for analyzing them.



Carnegie Mellon

Which actors are really critical depends on network topology e.g., its structure (organizational design)

	hierarchy	random	cellular	scale free
High degree centrality	☑			☑
High betweenness				☑
High cognitive demand			☑	
High knowledge exclusivity			☑	
High task exclusivity	☑		☑	☑

CASOS

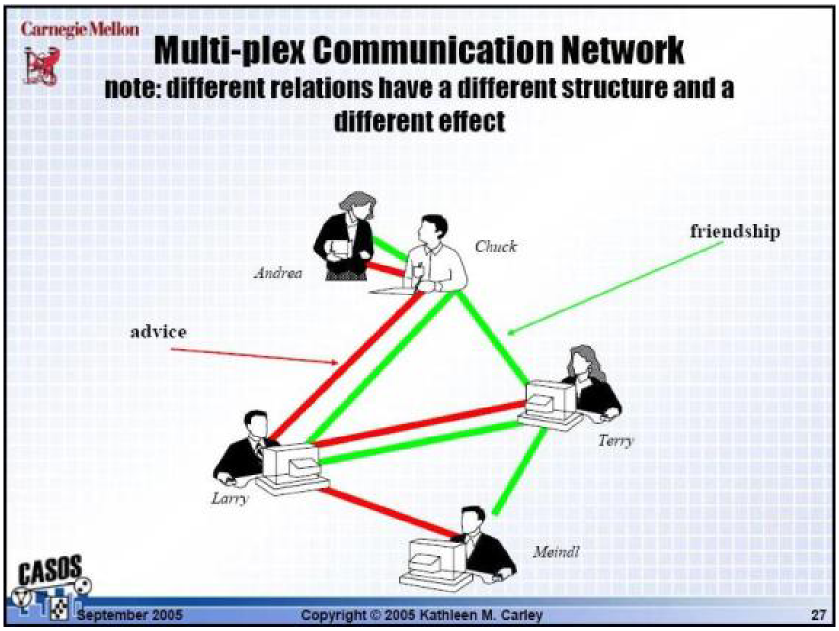
September 2005 Copyright © 2005 Kathleen M. Carley 38

Figure C-4: Different Kinds of Network Demand Different Measures.

Looking ahead, we will introduce exclusivities, cognitive demand and other measures further below in discussing multiplex and multimode measures. The point exemplified here by the measure of betweenness and degree centrality which we've already defined is that graph/network structure (topology) matters in the suitability of measures to be applied.

C.1.7 Measuring and Representing Multiplex Graphs/Networks

A multiplex graph/network is defined by more than one type of link/tie relationship among its nodes, e.g. friendship, advice, as illustrated in Figure C-5. Social environments exhibit multiplex graph/network connectivity.



Carnegie Mellon

Multi-plex Communication Network

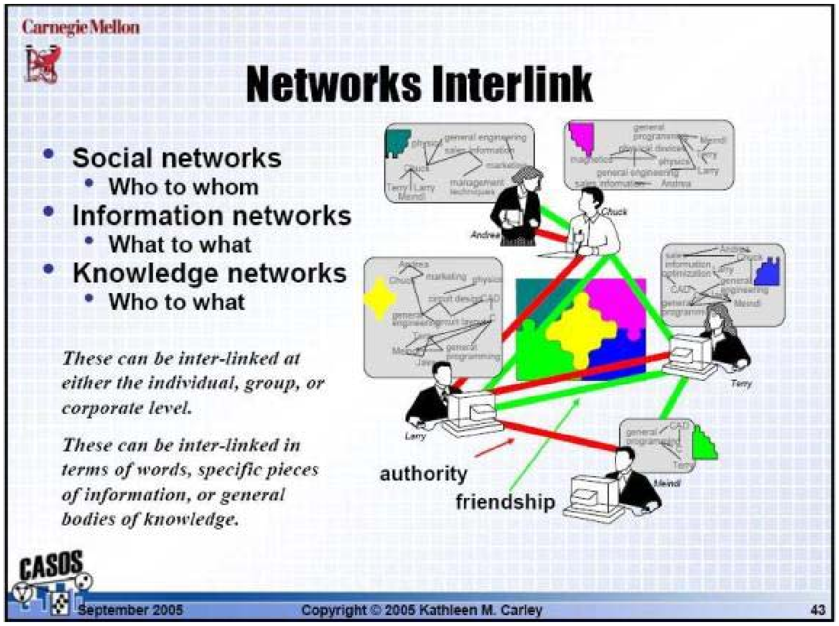
note: different relations have a different structure and a different effect

September 2005 Copyright © 2005 Kathleen M. Carley 27

Figure C-5: Different Kinds of Link Imply Multiple Independent Adjacency Matrices in a Multiplex Network.

Multiplex networks have multiple adjacency matrices, one for each type of link/tie relationship to be indicated. Figure C-5 shows how link/tie relationship types can be distinguished visually to good effect by displayed link colour.

A multimodal graph/network is defined by the existence within it of more than one type of node. Figure C-6 raises the question how a graph/network of nodes of one type is related to other graphs/networks of nodes of other types.



Carnegie Mellon

Networks Interlink

- Social networks
  - Who to whom
- Information networks
  - What to what
- Knowledge networks
  - Who to what

These can be inter-linked at either the individual, group, or corporate level.

These can be inter-linked in terms of words, specific pieces of information, or general bodies of knowledge.

September 2005 Copyright © 2005 Kathleen M. Carley 43

Figure C-6: Not only the Links, but also the Nodes, may be of Different Kinds in a Multimodal Network.

C.1.8 Measuring and Representing Multimodal Graphs/Networks

As suggested in Figure C-7, the mathematical technique for managing multiple modes (node types) in the adjacency matrix representation of a graph/network is to confine the indices for members of different modes (e.g. doctors and lawyers) to different non-overlapping regions of the adjacency matrix's enumeration of all nodes.

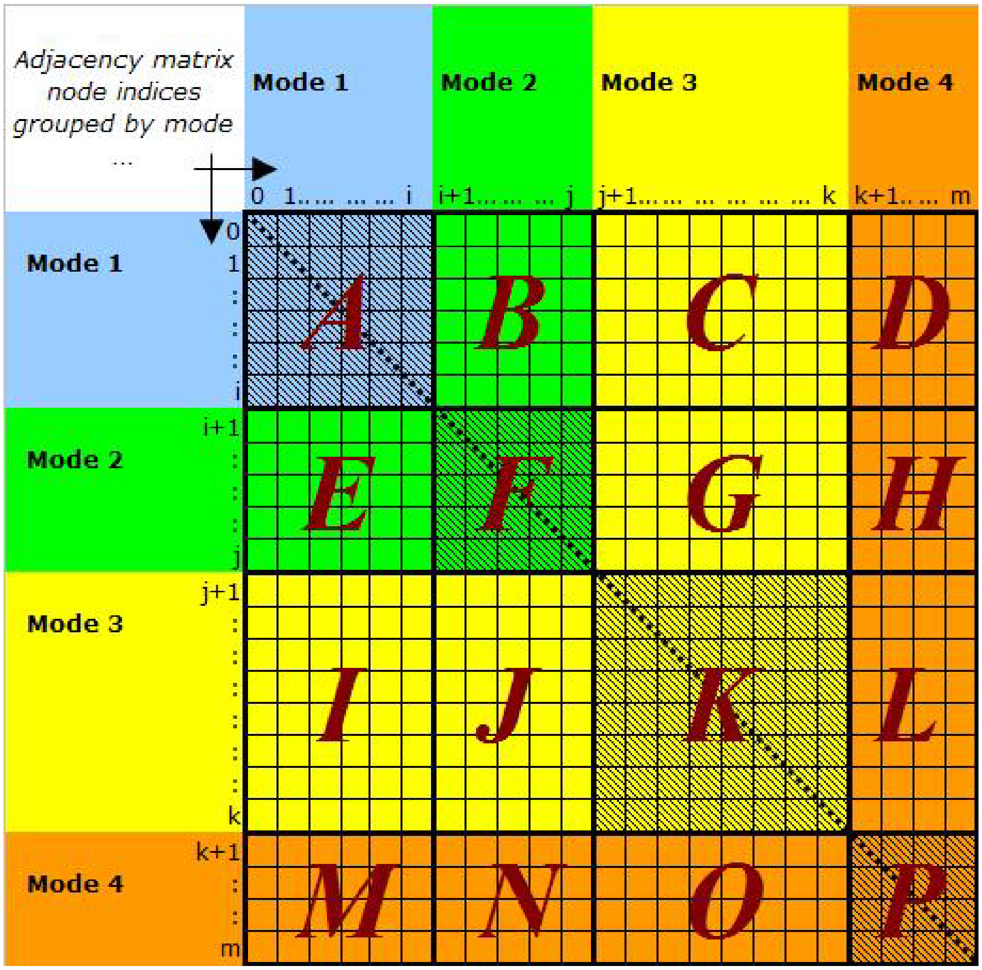
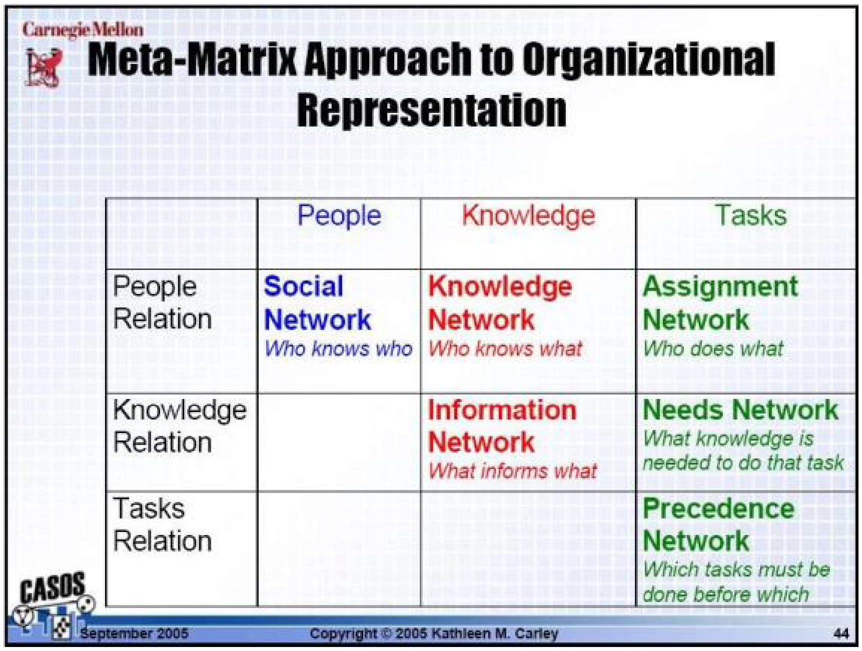


Figure C-7: Adjacency Matrix for a Multimodal Network, with Like-to-Like Links Collected Together so that they Appear in Square Regions around the Main Diagonal.

This way, the indication of ties among nodes of the same type (e.g. doctor-to-doctor or lawyer-to-lawyer "collaboration" relationships) are confined to square regions (same indexes in both dimensions) about the diagonal of the adjacency matrix, whereas the indication for ties among nodes of different types (e.g. doctor-to-lawyer "expert witness for" or lawyer-to-doctor "defends against malpractice claims" relationships) are "off-diagonal" and confined to rectangular (i.e. having possibly different index sizes and overall enumeration sub-ranges) regions away from the diagonal.

Figure C-8 lists examples of networks in which link/tie relationships (diagonal) are confined to the same mode and others (off-diagonal) in which they cross modes as illustrated in Figure C-7.



Carnegie Mellon

Meta-Matrix Approach to Organizational Representation

	People	Knowledge	Tasks
People Relation	**Social Network** *Who knows who*	**Knowledge Network** *Who knows what*	**Assignment Network** *Who does what*
Knowledge Relation		**Information Network** *What informs what*	**Needs Network** *What knowledge is needed to do that task*
Tasks Relation			**Precedence Network** *Which tasks must be done before which*

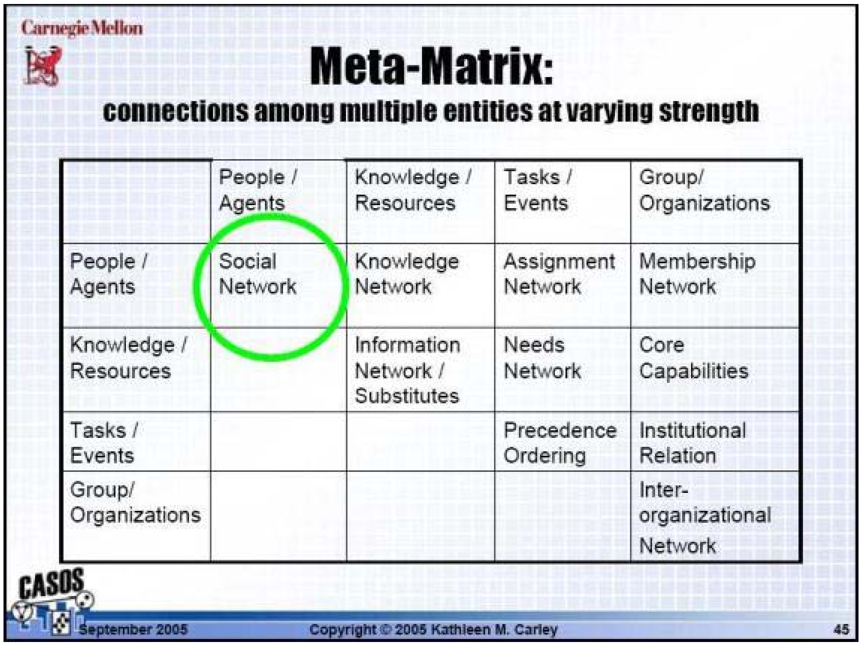
September 2005

Figure C-8: An Example of Multimodal Networks.

By abstracting the entire sub-range assigned to a mode into a single index, we obtain a metamatrix that neatly summarizes useful distinctions among the types of networks encoded in the adjacency matrix. This amounts to reducing the adjacency matrix for multimode graphs/networks in Figure C-7 to a 4 x 4 mode-to-mode matrix. The metamatrix is used to distinguish different graph/network measures appropriate to each graph/network region depending on the combination of modes involved as shown below.

Equally importantly, the metamatrix approach can be used to reveal the different kinds high-level information or analytic products that can be expected from analysis and measurement of the different kinds of networks it exhibits.

Figure C-9 reveals how traditional social network analysis was confined to analysis of single-mode graphs/ networks of actors. The other single-mode network analyses shown are familiar to other fields such as operations research while the cross-mode network analyses shown are part of the current evolution of social network analysis into a deeper interdisciplinary study of the social environment.



Carnegie Mellon

Meta-Matrix:

connections among multiple entities at varying strength

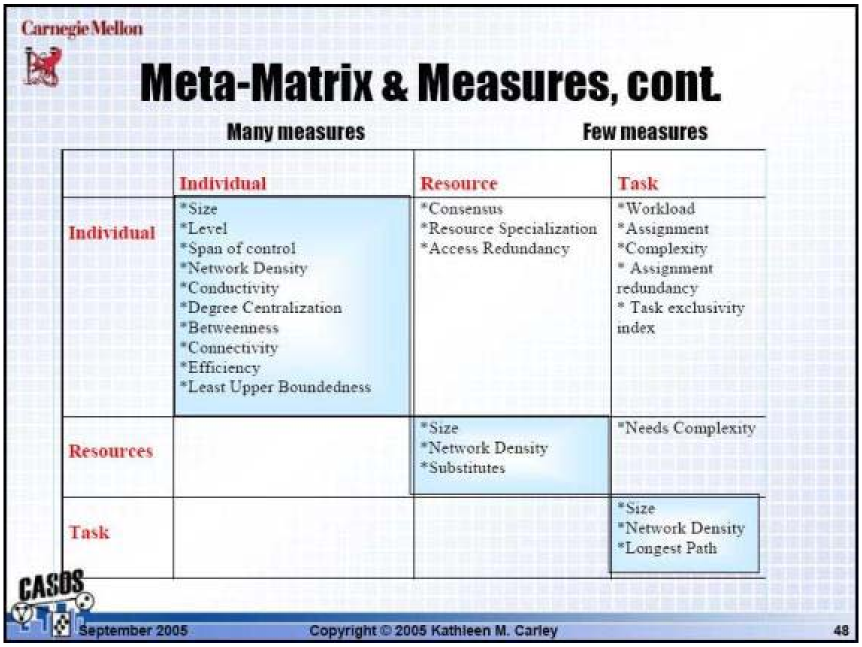
	People / Agents	Knowledge / Resources	Tasks / Events	Group/ Organizations
People / Agents	Social Network	Knowledge Network	Assignment Network	Membership Network
Knowledge / Resources		Information Network / Substitutes	Needs Network	Core Capabilities
Tasks / Events			Precedence Ordering	Institutional Relation
Group/ Organizations				Inter-organizational Network

CASOS

September 2005 Copyright © 2005 Kathleen M. Carley 45

Figure C-9: Different Kinds of Information can be Obtained from the Different Regions of a Metamatrix of Adjacencies.

Figure C-10 lists the more important graph/network measures to be applied in these different analyses. The next section discusses in greater detail the analysis and measurement of some of the multimode, off-diagonal networks mentioned in these figures.



Carnegie Mellon

Meta-Matrix & Measures, cont.

Many measures Few measures

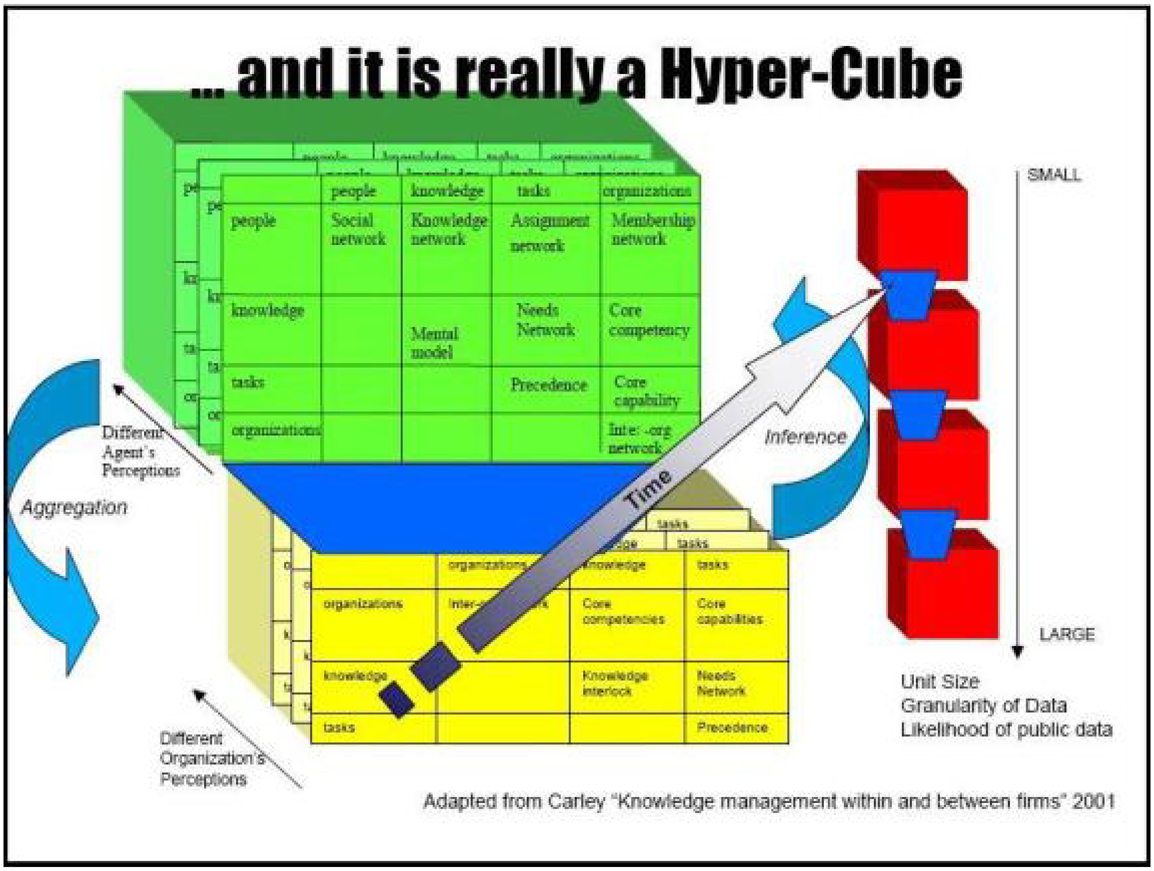
	Individual	Resource	Task
Individual	*Size *Level *Span of control *Network Density *Conductivity *Degree Centralization *Betweenness *Connectivity *Efficiency *Least Upper Boundedness	*Consensus *Resource Specialization *Access Redundancy	*Workload *Assignment *Complexity * Assignment redundancy * Task exclusivity index
Resources		*Size *Network Density *Substitutes	*Needs Complexity
Task			*Size *Network Density *Longest Path

CASOS

September 2005 Copyright © 2005 Kathleen M. Carley 48

Figure C-10: Some Measures and their Realms of Usefulness.

We summarize with Figure C-11, which shows (notionally) how the details of multiplex and multimodal graph/network metamatrices combine into higher (hyper-cubic) dimensions so as to account for more of the relationships important to a social environment.

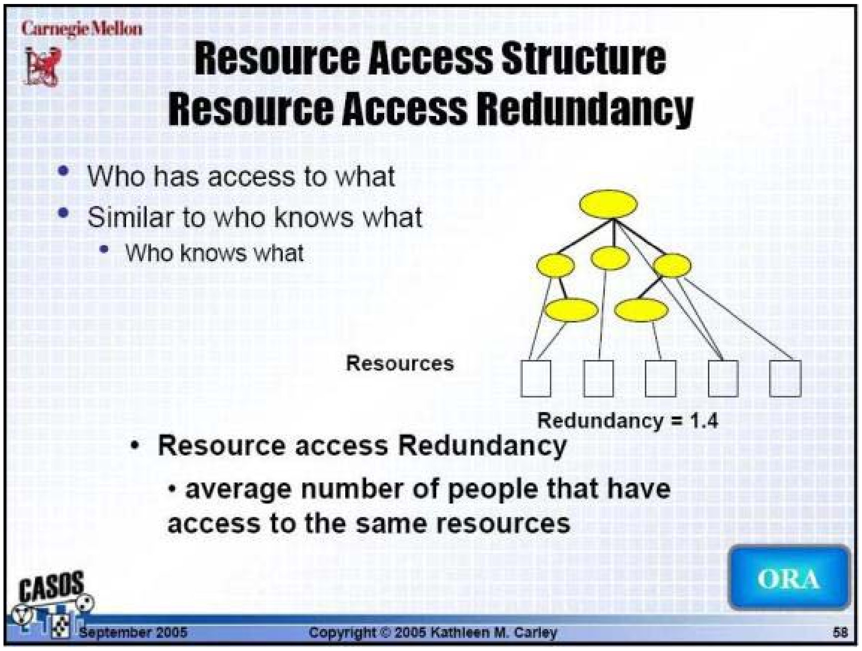


Adapted from Carley "Knowledge management within and between firms" 2001

Figure C-11: Schematic Suggestion as to the Way many Metamatrices may Combine in Several Additional Dimensions.

C.1.9 Cross-Mode Graph/Network Analysis and Measurement

The cross-mode resource access relationships shown in Figure C-12 tie actors to resources as well as to other actors.



Carnegie Mellon

Resource Access Structure
Resource Access Redundancy

- Who has access to what
- Similar to who knows what
  - Who knows what

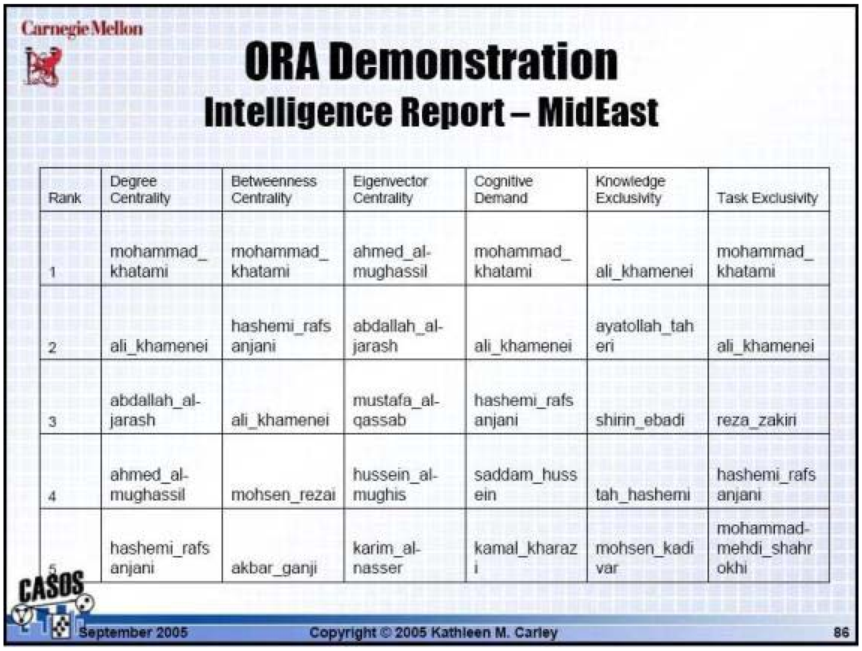
- Resource access Redundancy
  - average number of people that have access to the same resources

September 2005 Copyright © 2005 Kathleen M. Carley 58

Figure C-12: Resource Access; Tying Actors to Resources.

For resource nodes, *r*, the measure **Resource Access** defines the average number of people who have access to *r*, whence we define **Resource Access Redundancy** as the average resource access over all resources.

Figure C-13 ranks actors in a terrorist network according to some of the measures so far discussed. Importantly, the different rankings among the measures shown provoke questions essential to a more complete understanding of the essential interactions among organization members, especially members in lower ranks.



Carnegie Mellon

ORA Demonstration

Intelligence Report – MidEast

Rank	Degree Centrality	Betweenness Centrality	Eigenvector Centrality	Cognitive Demand	Knowledge Exclusivity	Task Exclusivity
1	mohammad_khatami	mohammad_khatami	ahmed_al-mughassil	mohammad_khatami	ali_khamenei	mohammad_khatami
2	ali_khamenei	hashemi_rafsanjani	abdallah_al-jarash	ali_khamenei	ayatollah_taheri	ali_khamenei
3	abdallah_al-jarash	ali_khamenei	mustafa_al-qassab	hashemi_rafsanjani	shirin_ebadi	reza_zakiri
4	ahmed_al-mughassil	mohsen_rezai	hussein_al-mughis	saddam_hussein	tah_hashemi	hashemi_rafsanjani
5	hashemi_rafsanjani	akbar_ganji	karim_al-nasser	kamal_kharazi	mohsen_kadivar	mohammad-mehdi_shahrokhi

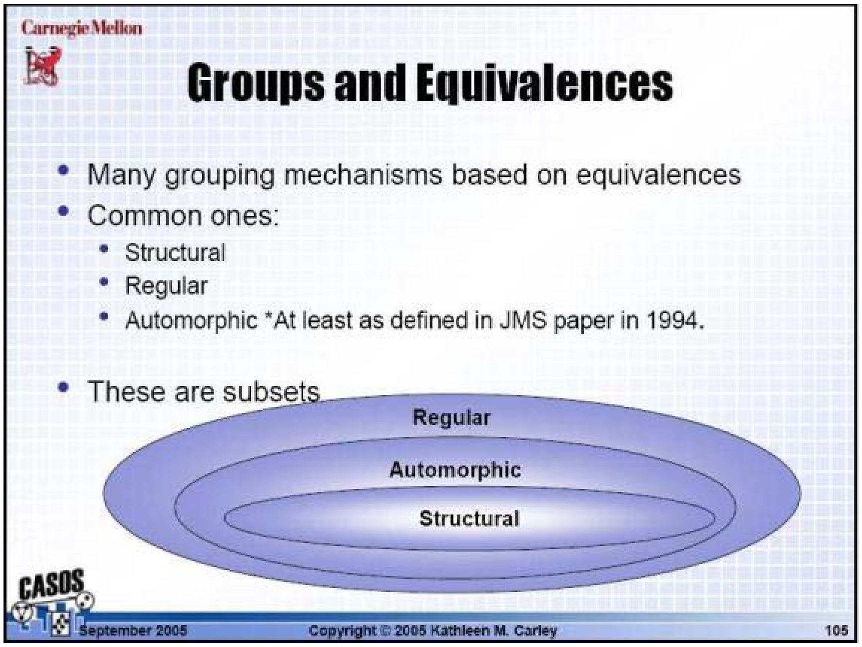
CASOS

September 2005 Copyright © 2005 Kathleen M. Carley 86

Figure C-13: Rankings of Some Terrorists on Several Measures.

C.1.10 Groups, Equivalences and Colorations

Figure C-14 introduces the notion of a group in term of various types of equivalence between nodes in a graph/ network.



Carnegie Mellon

Groups and Equivalences

- Many grouping mechanisms based on equivalences
- Common ones:
  - Structural
  - Regular
  - Automorphic *At least as defined in JMS paper in 1994.
- These are subsets

CASOS

September 2005 Copyright © 2005 Kathleen M. Carley 105

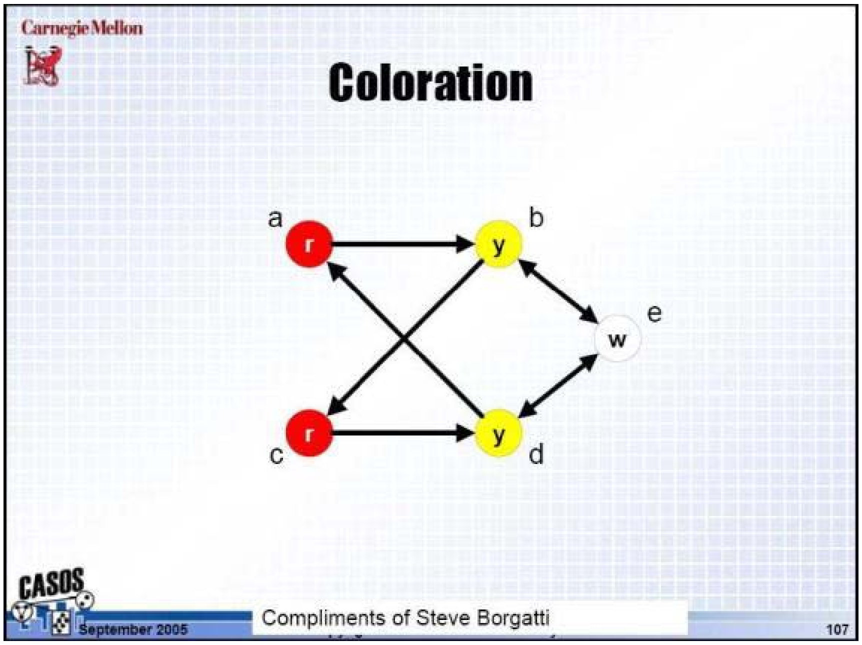
Figure C-14: Types of Grouping of Nodes or Sub-Nets.

Groups as equivalence classes satisfy the following three conditions on membership according to their defining equivalence relation, E:

- Transitivity - $(a,b), (b,c) \in E \rightarrow (a,c) \in E$
- Symmetry - $(a,b) \in E \leftrightarrow (b,a) \in E$
- Reflexivity - $(a,a) \in E$

A colouration is just a partition of the nodes in a graph/network, i.e. an assignment of the nodes to mutually exclusive and exhaustive classes according to equivalence relations. The colour, $C(v)$, of a node, v, is then just the equivalence class to which it belongs.

Groups are distinguished by colourations according to their equivalence classes. In Figure C-15, red nodes have in-links only from and out-links only to yellow nodes; yellow nodes have in-links from and out-links to red nodes, as well as bidirectional links to the only white node; and the white node has only bidirectional links to yellow nodes. The equivalences are in the link structure, and the relationship of the red and yellow nodes makes this a striped multimodal network.



Coloration

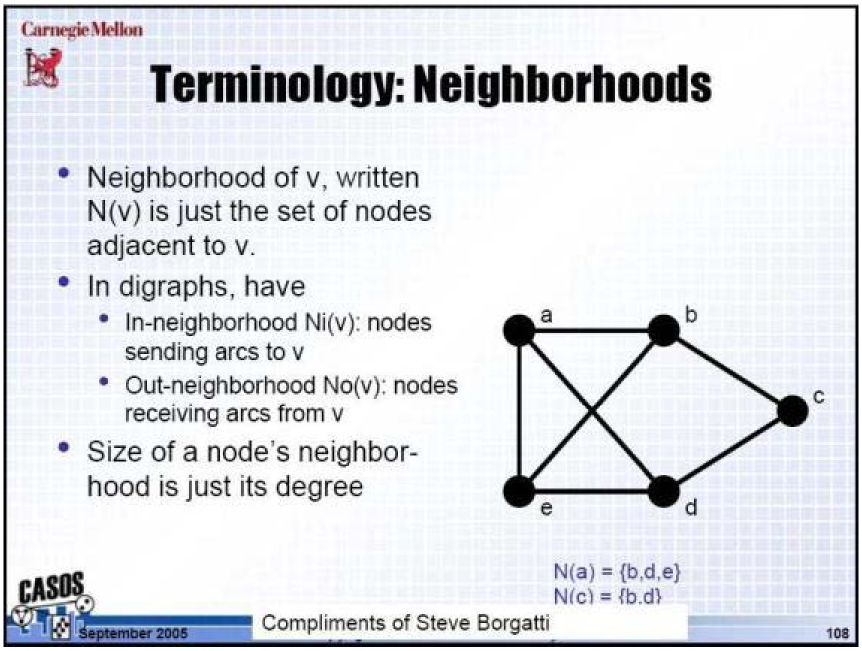
September 2005

Compliments of Steve Borgatti

107

Figure C-15: Coloured Nodes of Three Different Equivalence Classes.

The neighbourhood of a node, v, in a network is defined in Figure C-16. It is the union of the set of nodes (in-neighbours) sending an arc to node v and those (out-neighbours) receiving an arc from v.



Carnegie Mellon

Terminology: Neighborhoods

- Neighborhood of v, written N(v) is just the set of nodes adjacent to v.
- In digraphs, have
  - In-neighborhood Ni(v): nodes sending arcs to v
  - Out-neighborhood No(v): nodes receiving arcs from v
- Size of a node's neighbor-hood is just its degree

N(a) = {b,d,e}
N(c) = {b,d}

CASOS

September 2005

Compliments of Steve Borgatti

108

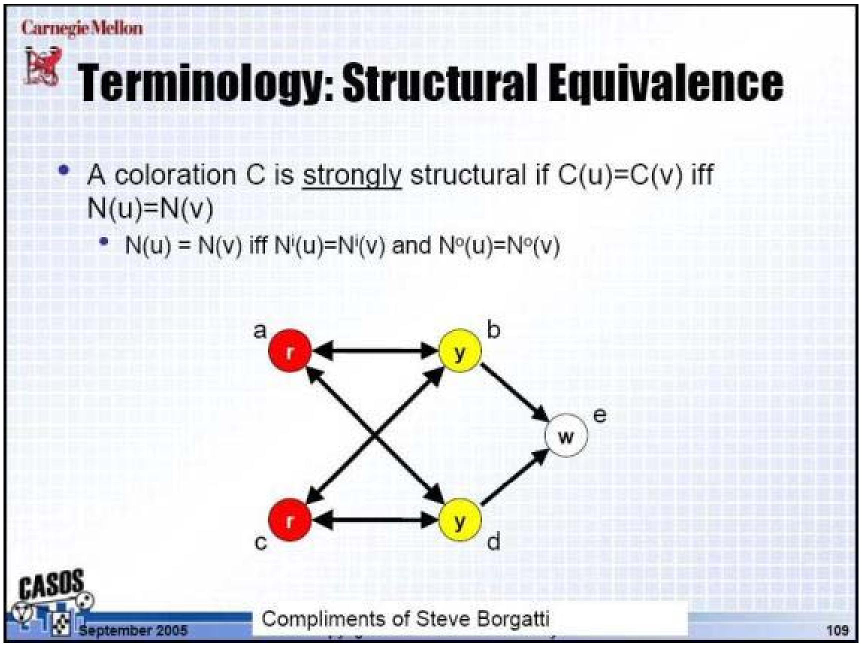
Figure C-16: Neighbourhood of a Node.

C.1.11 Types of Equivalence

In this section, we discuss in more detail the notions of structural, automorphic and regular equivalence mentioned above. Each type of equivalence imposes progressively stronger conditions on equivalence.

Structurally equivalent nodes have the same degree and belong to the same cliques. They are distinguishable only by label and are, therefore, said to be perfectly substitutable in the social environment (e.g. same contacts, resources).

Strongly structural colourations (equivalence classifications) are those in which nodes of the same colour have the same neighbourhoods as shown in Figure C-17. Viewed as actors in a social network, structurally equivalent nodes face the same social environment. There are similar forces affecting them. They are subject to the same influences. On average, they hear things equally early, are influenced similarly and have similar things to cope with. Structural equivalence is used to capture and model inter-node relationships.



Carnegie Mellon

Terminology: Structural Equivalence

- A coloration C is <u>strongly</u> structural if C(u)=C(v) iff N(u)=N(v)
  - N(u) = N(v) iff $N^i(u)=N^i(v)$ and $N^o(u)=N^o(v)$

CASOS

September 2005

Compliments of Steve Borgatti

Figure C-17: The Nodes of Different Colours are Structurally Equivalent.

Mechanisms can often be revealed as structural equivalences among nodes. Structural equivalence does not quite suit the notion of a social role, however, because it will often over-determine a given social role that in fact crosses many colouration pairs.

Structural equivalence is computed as a similarity or distance measure between rows of an adjacency matrix using correlation, Euclidean distance, etc. Diagonals of adjacency matrices of course represent identities which, trivially, are also similarities on any such similarity measure. It is often useful to eliminate these trivial similarities in the interests of brevity and visual simplification. A proximity matrix encoding the similarity measure among nodes in a network can then be computed bottom up using a clustering or minimum-distance-search (MDS) algorithm. The problem with this bottom-up approach is the stopping condition on the clustering or MDS criterion. Conversely, other tools use a correlation algorithm iteratively in a *top-down* fashion and do not have a stopping condition problem. However, the top-down correlation method suffers from the defect that it imposes structure *a priori* by choice of correlation criteria.

To address the shortcomings of structural equivalence, we consider another key grouping method – isomorphisms among (sub)graphs. Here the mapping constituting the graph isomorphism preserves adjacency structure. The analytical property of preserving adjacency structure can be as straightforward as one-to-one mappings between graphs. More subtle examples involving more complex graphs/networks can be easily constructed. But in these cases, the reliance on algorithms to determine preservation of adjacency structure is called for, there being no obvious intuitive visual test that will capture every case infallibly.

More important, an automorphism is an isomorphism of a graph onto itself as indicated by the mapping, $p(v)$, among nodes, v, of the graph. Automorphic mappings of simple graphs are also easily defined. These ideas lead to the key notion of *automorphic equivalence* which better supports our main intuitions about groupings. Automorphic equivalence is truly structural and positional and is not confounded by contiguity the way ordinary structural equivalence is. Moreover, as a generalization of structural equivalence, automorphic equivalence accounts for additional mappings among sub-graphs that substantiate the essentials of the *social role* concept.

C.2 REFERENCES

[1] Carley, K.M., "Dynamic Network Analysis", *Short Course*, Center for Computational Analysis of Social and Organizational Systems, September 2005, Available from: http://www.casos.cs.cmu.edu/.

[2] Memon, N. and Larsen, H.L., "Practical Algorithms of Destabilizing Terrorist Networks", In *Proceedings of IEEE Intelligence Security Conference*, San Diego, Number 3976: In *Lecture Notes in Computer Science*, pp. 398-411, IEEE Intelligence Security Conference, Springer-Verlag, January 7, 2006.

Annex D – ASPECTS OF THE APPLICATION OF INFORMATION THEORY TO VISUAL COMMUNICATION

J.T. Bjørke

Information theory can play an important role in understanding how to design for efficient coding of visual communication. This proposition will be elaborated upon in this Annex.

D.1 INTRODUCTION

Recent research has demonstrated that the mathematical theory of communication [31] can be applied to optimize the design of visual images like maps and images for scientific communication. Although there is still a lack of comprehensive theory on how to quantify the efficiency of visual communication, information theory can provide a mathematical basis for better understanding of some aspects of this type of communication.

Information theory was introduced into cartographic research in the 1970s, but at the time it met critical voices and did not lead to new algorithms for automated map design. From the end of the 1980s cartographic algorithms inspired by information theory can be seen in the literature.

Moles [23] points out that the most obvious failure of the theory in its simplest form is that it appears an atomistic theory which tends to explain reality by decomposing it into simple elements. Head [14] claims that it is not fully understood how to quantify the information itself.

> "*It came to be recognized, however, that map readers often seemed to get things from map reading that were not consciously designed-in by the cartographer, and this made measurement of information loss a fuzzy business*".

Neumann [25] comments the criticism of the 1970s:

> *"The communication concept had one weak point – the use of information theory was mechanically conditioned by the application of Shannon's theory of communication. Consequently, it was criticized by Salichtchev [30], Robinson and Petchenik [28], and other authors in the 1970s. The critics were particular to point out that the conventional process of communication, accompanied with losses in transmitted information, could not be used as a model of the cartographic process which, in contrast, produced an increase in the amount of information."*

The type of criticism cited, actually demonstrates the limitations of Shannon information theory and the problems that arise when attempting to apply it to map evaluation or design of scientific images. The earlier criticism must be seen in the light of the development of computers. In the 1970s cartographic theory was highly related to manual map production methods. Information theory offers methods to optimize the content of maps and can be computationally expensive. The role of information theory today is more to inspire the design of computer algorithms for zoom in and out of a spatial data set than to offer a comprehensive theory for visual communication.

Robinson and Petchenik [28] correctly point out that the positional factor of a map must be considered if information theory is to be applied to cartography. Since the cartographic application of information theory of the 1960s and 1970s did not emphasize the positional component of a map, information theory was probably brought into discredit.

Knöpfli [18] explains the difference between aerial photos and maps in terms of information theory and shows that the amount of information in aerial photos as well as maps, can be reduced by misinterpretations of the relevant messages. He nicely demonstrates the effect of distorted (noisy) information transmission and sets up two steps in order to reduce the loss of relevant information:

1) Omit the irrelevant characteristics; and
2) Strengthen the relevant characteristics.

These rules can be reformulated to:

1) Not overloading the map with information (in this context information has the narrow meaning as syntactic information); and
2) Maintaining a sufficient "visual distance" between the map symbols to make them distinguishable (in this context "visual distance" can be Euclidian distance in the map plane or distance defined in the domain of the visual variables such as colour, shape and size).

Even if these rules are simplistic and general, they are very important to consider in map design and scientific visual communication.

Communication can be divided into the three levels:

- **Syntactic** regarding the relationship among the signs that are employed in the communication;
- **Semantic** regarding the relationship between the signs and the entities which they represent, that is, the designation of the meaning of the signs; and
- **Pragmatic** regarding the relationship between the signs and their application.

Shannon and Weaver [31] are distinct about these aspects of communication. In their terminology the three aspects are termed *levels of communication problems* and are given the abbreviations: Level A, Level B and Level C; which relate to the syntactic, semantic and pragmatic aspects respectively. Shannon and Weaver emphasize that at Level A they use the word *information* in a special sense that not must be confused with its ordinary usage. In particular, information must not be confused with meaning. To be somewhat more definite, the amount of information is defined, in the simplest cases, to be measured by the logarithm of the number of available choices.

It may happen that some of the earlier criticism of the application of information theory would have been moderated if one were more distinct about the three levels of communication problems and evaluated the relevance of information theory specifically to each of the levels.

D.2 PREVIOUS ATTEMPTS AT APPLYING COMMUNICATION THEORY TO CARTOGRAPHY

Information theory introduces measures of variation. Therefore, the application of the theory to the modelling of visual communication requires that the statistical properties of the information source is well understood and the important characteristics of the communication are described as constraints, weight functions, statistical measures, etc.

Sukhov [33] proposes an atomistic method to compute the entropy of a map. This is based on a method which breaks a map into discrete elements. A statistical sampling method is used for selecting typical unit areas from the map for measuring the entropy. The method is applied to different sub-systems of the map, i.e. different themes such as hydrography, relief and roads. Finally, the map entropy is computed as the sum of the entropies of its different sub-systems. Sukhov is distinct about the significance of the correlation between the sub-systems. Since the sub-systems of the study were weakly correlated, it gave Sukhov the basis for using the joint entropy computation (see Equation (18) in the Appendix). Sukhov's contribution gives insight into the significance of correlation in the computation of the joint entropy of different information sources.

Two papers by Knöpfli, [17] and [18], explain some features of cartographic generalization in terms of Shannon entropy. The first [17], demonstrates that some information can be derived from the structure of, what is termed the embedding space, using inductive reasoning. For example, if a city is located on both sides of a river, we can conjecture that there must be a bridge between the two parts of the city. Since the no-bridge case would be very unusual, its information value is very high. Therefore, the information that there is a bridge has a lower information value than the information that there is no bridge. The example demonstrates that spatial correlation and spatial context should be considered in the entropy computations. In the second paper [18], the difference between aerial photographs and maps are explained in terms of information theory. The paper demonstrates very clearly that the scatter of the relevant messages (noise) leads to loss of information.

> *"It is always claimed that aerial photographs contain much more information than maps. Since I have dealt with the production of topographic maps from aerial photos for years, I am familiar with the advantages and disadvantages of both products and have never agreed with this assertion."*

Bjørke and Aasgaard [9] propose information theory as a part of the concept of what they call "cartographic zoom". This is a real time concept which aims to generate map versions adjusted to the dynamic change of map scale on a computer screen. Information theory is described as a tool to measure the amount of information on a map and this is proposed to be integrated into a sub-system which controls the number of map symbols and their visibility. They emphasize that they use the term information in a narrow sense, and that their use of "information" has no connection with "meaning". Therefore, their application of information theory is restricted to the syntactic level of information, i.e. level A according to the terminology of Shannon and Weaver.

Bjørke [5] demonstrates how information theory can be used to control the generalization process in the two cases:

1) The selection of the number of classes in choropleth raster maps; and
2) The selection of parameter values in automated line generalization.

In both cases the channel capacity of the maps was computed. In the first case the borders between the raster elements (pixels) were selected as events for the entropy computation. An investigation of some subjects gave the probabilities that the different grey values were misinterpreted. Then the channel capacity of a random and a correlated choropleth map were computed. From this computation, an optimum number of classes was derived for the two maps. In the second case, the angular change of the line to be generalized served as a basis for the entropy computation. Based on a model of the minimum separable distance between the events, an optimum value of the line generalization parameter was derived. Bjørke and Midtbø [10] go further and apply information theory to contouring from digital elevation models. In this case the underlying terrain model was simplified, not the contour lines themselves, and an optimum generalization parameter value was derived. This paper also loosely proposes an information theory method to compute an optimum contour interval.

Bjørke [2] proposes a framework for the application of information theory to cartographic map design. He introduces the concept of different types of entropies in a map and proposes a model for map design based on information theory. At the same time Neumann [25] presents a paper where the topological entropy of a map is focused. The topological entropy of Neumann is computed from dual graphs (Region Adjacency Graphs). Bjørke [2] also defines a topological entropy, but the entropies of Neumann and Bjørke are different. Further applications of information theory to map design are reported by [11] (point symbol maps), [20] (map complexity), [3] (road maps) and [4] (road maps).

D.3 APPLICATION OF INFORMATION THEORY TO NEUROBIOLOGY

Resent work in neurobiology, opens the perspective that the quantitative methods offered by information theory can be utilized in understanding the whole process from creating the signs on a visual display to the neurobiological processes that transmits the input signals to the brain's cognition of an image. Moreover, the neurobiological research can inspire design of algorithms for scientific visualisation.

Simoncelli [32] deals with how the Efficient Coding Hypothesis can help understanding properties of visual systems. The author shows that this hypothesis has led to studying the influence of environmental statistics on neural response. Simoncelli points out there are difficulties in the application of information theoretic modelling to neurobiological systems. For example, difficulties lie in the definition of the input (what is a natural image) and the output (how to define the neural response) as well as in incorporating realistic constraints (e.g. noise and metabolic costs) and computational goals. Therefore, the application of information theory to understanding visual systems is not straight forward, but despite these challenges researchers in the neurobiological field has recognized the quantitative strengths of the theory.

From experiments Reynolds [26] concludes that spatial attention causes changes in the neuronal responses that are similar to the effects of increasing the effective contrast of the attended stimulus. In a similar kind of research Kastner [15] investigates brain areas associated to attention filters. In this study it was found that certain brain areas appear to be important sites at which attention filters out unwanted information by means of receptive field mechanisms. Despite what the two authors cited, they did not use entropy measures in their papers considered, their terminology is interesting from an information theoretic point of view. For example, they use terms like increasing the effective contrast and filter out unwanted information.

D.4 QUANTIFYING THE INFORMATION CONTENT OF A MAP

D.4.1 Spatial Correlation

When applying Shannon information theory in cartography, we face the problem of how to deal with spatial correlation. An aspect of the problem is demonstrated in Figure D-1.

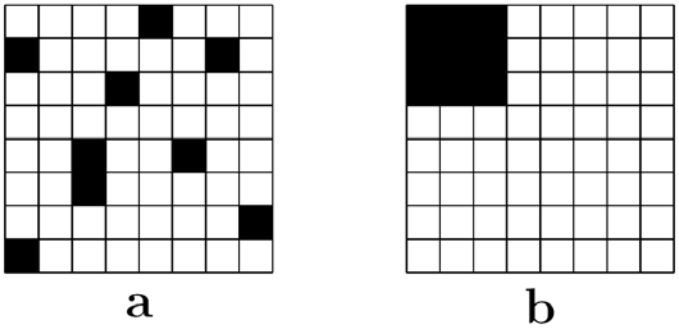


Figure D-1: Different Patterns of Binary Images. The patterns made by the two images are different, but the number of black pixels is equal in both.

Both the images in the figure consist of 9 black and 55 white pixels, but the patterns in the two images are different. If we calculate the entropy of the two images on basis of counting the number of black and white pixels, the entropy is computed as:

$$H_a(X) = H_b(X) = -\frac{9}{64} \cdot \log_2 \frac{9}{64} - \frac{55}{64} \cdot \log_2 \frac{55}{64} = 0.586.$$

where 9/64 is the probability of finding black pixels and 55/64 is the probability of finding white pixels. To an observer the pattern of image (b) looks more ordered than the pattern of image (a), but we have computed identical values for their entropies. The reason is that the events in the message are spatially correlated and we have not modelled that correlation. The spatial correlation between neighbouring pixels of an image can be taken care of by replacing the values of the pixels by their differences. Based on this idea, the previous entropy computation will be reformulated. If two neighbouring pixels have the same colour, we define their difference to be positive. Otherwise, if the pixels have different colours, their difference is defined as negative. According to this strategy, the entropy of a binary image can be defined as:

$$H(X) = -p^+ \cdot \log_2 p^+ - p^- \cdot \log_2 p^- \tag{1}$$

where p^+ is the probability of $(black, black)$ and $(white, white)$ neighbours while p^- is the probability of $(black, white)$ and $(white, black)$ neighbours. Applying this technique to the images of Figure D-1, we get $H_a(X) = 0.825$ and $H_b(X) = 0.301$. Image (b) now has lower entropy than image (a) which puts the images into a sequence corresponding to our visual judgment.

Gatrell [12] proposes computing the entropy of a binary image as a weighted mean value of the entropy at the different orders of neighbourhood. The computation can be done by applying Equation (1) to the different orders of neighbourhood. We can set up the equation:

$$H(X) = \sum_{k=0}^{n} w(k) \cdot H(X)_k \tag{2}$$

where $w(k)$ is a weight function and k is the order of neighbourhood. Equation (2) has some conformity with the joint entropy in Equation (18) (Appendix). If $w(1) = w(2) = \cdots = w(k) = 1$ and the different levels are independent in the probabilistic sense, the equation corresponds to the joint entropy of the k information sources. The weight function is used to control the size of the neighbourhood to be evaluated. A high value of k corresponds to a global neighbourhood while a small value corresponds to a local neighbourhood.

D.5 MAP INFORMATION SOURCES

When applying information theory to cartography, we should carefully identify the elements which make up the variation of a map. As earlier stated, this Annex mainly deals with communication problems at the syntactic level of cartographic communication. Therefore, our identification of information sources only concerns the syntactic properties of the map. For the following discussion we need a definition of the terms map entity and map information source.

Definition 1 – *A map entity can be a map symbol, a part of a map symbol, groups of map symbols, an attribute of a map symbol or a derived characteristic of a map which can serve as an entity for entropy computations.*

Definition 2 – *A map information source* X *is an object which contains a set* X *of map entities and a characteristic* C *of them which make up their variation.*

The visual variables identified by Bertin [1], will serve as a basis for our classification of map information sources. According to Bertin, the variables which are used to manipulate the map symbols are: X, Y (the two dimensions of the plane), size, value, texture, colour, orientation and shape. Bertin operates with two components of the map plane, the X and Y co-ordinates, as visual variables. In entropy computations it is more appropriate to distinguish between three components of the map plane as illustrated in Figure D-2.

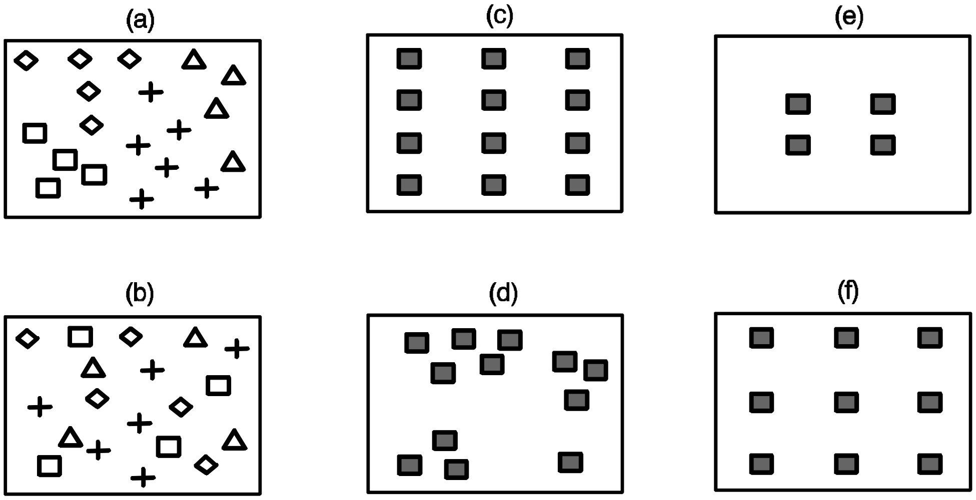


Figure D-2: Entropies of the Map Plane. Image (a) and image (b) demonstrate topological entropy, image (c) and (d) demonstrate the concept of metrical entropy whereas image (e) and (f) demonstrate positional entropy.

A first entropy is derived from images (a) and (b) in Figure D-2. From a visual point of view it is clear that in Figure D-2, map (a) is more ordered than map (b), but the number of different map symbols and the (X, Y)

positions occupied by the set of symbols are equal in both the maps. The entropy of the kind considered in images (a) and (b), will be termed *topological* entropy.

Definition 3 – *The topological entropy of a map considers the topological arrangement of the map entities.*

A second entropy which can be derived from images (c) and (d) in Figure D-2, is the *metrical* entropy of a map.

Definition 4 – *The metrical entropy of a map considers the variation of the distance between the map entities.*

A third type of entropy which can be derived from images (e) and (f) in Figure D-2, is *positional* entropy.

Definition 5 – *The positional entropy of a map considers all the occurrences of the map entities as unique events. In the special case that all the map events are equally probable,* $H(X) = \log_2 n$ *; where* n *is the number of entities.*

The term positional entropy is motivated from its relation to the number of positions occupied by the map entities. If we assume that each map entity occupies one position, the positional entropy is simply computed from counting the number of map entities. Our definition of topological entropy and metrical entropy correspond to the definitions of [6] while the definition of positional entropy corresponds to the definition of density entropy.

The computation of topological entropy and metrical entropy of the point symbol maps in Figure D-2 requires a spatial concept. There may be several strategies which can be applied to this, but I will propose a method similar to that used for the binary image case in Figure D-1. Imagine a point symbol and some neighbouring symbols. We will define *the visual area* of a point symbol as its Thiessen polygon. Since a Delaunay triangulation is the dual of a set of Thiessen polygons [19], we will base the neighbourhood definition in a point symbol map on a Delaunay triangulation. This idea is demonstrated in Figure D-3.

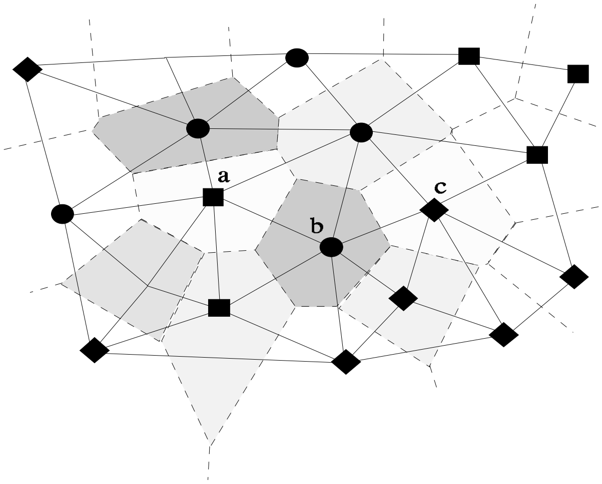


Figure D-3: A Point Symbol Map and its Thiessen Polygons.

A Thiessen polygon is constructed around each map symbol. Therefore, the map symbols are nodes in a network created by a Delaunay triangulation. Given two nodes in the network created by the Delaunay

]

triangulation, the order of the neighbourhood is computed by counting the number of edges on the shortest path, by the number of links, between the points considered. For example, point (b) is a 1st order neighbour of point (a) whereas point (c) is a 2nd order neighbour of point (a). Since we have a strategy to define neighbours, we can apply the difference technique of the binary image in Figure D-2. The topological entropy is based on computing the probability of different types of binary relations between the map symbols. In Figure D-2, for example, we get the set E of entities (relations):

$$E = \begin{bmatrix} E_{11} & E_{12} & E_{13} & E_{14} \\ E_{21} & E_{22} & E_{23} & E_{24} \\ E_{31} & E_{32} & E_{33} & E_{34} \\ E_{41} & E_{42} & E_{43} & E_{44} \end{bmatrix} \tag{3}$$

The definition of the entities in Equation (3) is more complete than the definitions in Equation (1), since in Equation (3) the symmetry $(black, white)$, $(white, black)$ and $(black, black)$, $(white, white)$ is regarded as distinct events. If the $0-$th order neighbourhood is only considered, we get the sub-set:

$$E_0 = \begin{bmatrix} E_1 & E_2 & E_3 & E_4 \end{bmatrix} \tag{4}$$

which corresponds to the selection of entities proposed in [18]. Applying the method considered to different orders of neighbourhood, we get a set of entropies. A mean value for the set can be computed as a weighted sum of the entropies at different orders of neighbourhood (Equation (2)). For the metrical entropy of the maps in Figure D-2, we can simply calculate the Euclidian distance between the neighbouring map symbols and apply the distance *differences* rather than the distance values themselves as entities. As for the topological entropy, the metrical entropy can also be computed at different orders of neighbourhood.

Equation (3) shows a relation between topological entropy and the visual variable shape. In this case the differential variable shape does distinguish between the 9 elements of set E. To be more definite, the visual variables: size, value, texture, colour, orientation and shape belong to the attribute domain of the map. A class name for a specific group of map information sources will be introduced.

Definition 6 – *Map information sources are orthogonal if none of the information sources can be derived from combining some of the other information sources.*

Definition 7 – *The topological, metrical and positional entropies have orthogonal map information sources; which are information sources of the spatial domain of a map.*

Definition 8 – *The visual variables as: size, value, texture, colour, orientation and shape have orthogonal map information sources; which are information sources of the attribute domain of a map.*

D.6 SIMILARITY GRADE, TRANSITION PROBABILITY AND EQUIVOCATION

If the map user is uncertain about the map symbols actually received, this uncertainty is defined as equivocation [31][18]. Knöpfli [18] clearly shows that the “visual distance” between the map symbols is important for the perception of the symbols, i.e. at a small visual distance there is a chance that one symbol is interpreted as another symbol. For example, if two lines A and B are very close to each other, it may be difficult to visually separate the one line from the other. Therefore, some parts of line A may be interpreted as line B. Another example is that if two symbols have similar colours, the colour of one symbol may be interpreted as the colour of the other symbol. If the map designer planned to distinguish between the two colours, the similarity in colour may cause confusion for the map reader. The perceived similarity between map symbols calls for a definition:

Definition 9 – *A function $\mu(x,y)$ which defines the grade of perceived similarity between two map entities x and y, will be termed similarity function. The similarity is measured on the interval $[0,1]$ of real numbers. If x and y are clearly separable, the similarity grade is 0. If x and y are completely unseparable, the similarity grade is 1.*

The computation of the similarity function for a particular map information source, is not a trivial task, because the perceived similarity between map entities may be influenced by several types of phenomena. For example: [13] shows that the perceived size of a circle may be biased by its map context and [27] point out that the perceived size of a line is influenced by its background colour. Methods to compute the similarity function and which perceptual phenomena to consider, are mainly outside the scope of this paper.

In equivocation computations we need to know the transition probabilities, i.e. the conditional probabilities in Equation (15) or (16) (Appendix). Similarity grade and transition probability are related to each other, but they are different. The difference will be explained and a mapping from similarity grade to transition probability will be proposed. Our definition of the similarity function corresponds to the definition of the membership function in fuzzy set theory (fuzzy set theory is explained in [16], for example). In fuzzy set theory the membership function assigns a value to the members of the set. The membership function by which a set A is defined, has the form:

$$\mu_A : X \to [0,1]$$

where $[0,1]$ denotes the interval of real numbers from 0 to 1, inclusive. The grade of membership of an element x in A is written as $\mu_A(x)$. Sometimes $\mu_A(x)$ is termed the possibility that x is a member of A. The concept of possibility and probability are both used to represent and manipulate imprecision or uncertainty. In everyday speech the terms possibility and probability are sometimes used interchangeably. However, there is a fundamental difference between possibility and probability. For example, the probabilities must sum to 1 whereas the possibility values are not restricted in such a way.

Consider a set X of map entities and a set Y of perceived map entities and the relation $R(X,Y)$ defined by the Cartesian product $X \times Y$. Formally, $X \times Y = \{(x,y) \mid x \in X \ \ and \ \ y \in Y\}$. Let every tuple (x,y) of the relation be assigned a similarity grade $\mu(x,y)$, i.e. the tuples may have varying degrees of membership within the relation. Since a map entity should be similar to itself by the maximum rate of similarity, the tuple (x,y) is given the membership grade 1 if $x = y$. From $R(X,Y)$ we select the similarity class $x \times Y = \{(x,y) \mid y \in Y\}$ and compute the class sum $\sum_{y \in Y} \mu(y \mid x)$. Our definition of similarity class has the properties of similarity class in fuzzy set theory. In fuzzy set theory a similarity class is defined as a fuzzy set in which the membership grade of any particular element represents the similarity of that element to the element x ([16], page 83). The mapping from grade of similarity to transition probabilities (conditional probabilities) can then be done for each similarity class as:

$$p(y \mid x) = \frac{\mu(y \mid x)}{\sum_{y \in Y} \mu(y \mid x)} \quad \text{for each } y \in Y \tag{5}$$

where the notation $\mu(y \mid x)$ is equivalent to the notation $\mu(x, y)$ and should be read as: the grade of membership of the relation from x to y. Applying Equation (5) to all the classes of X ensures that:

$$\sum_{y \in Y} p(y \mid x) = 1 \text{ for each } x \in X .$$

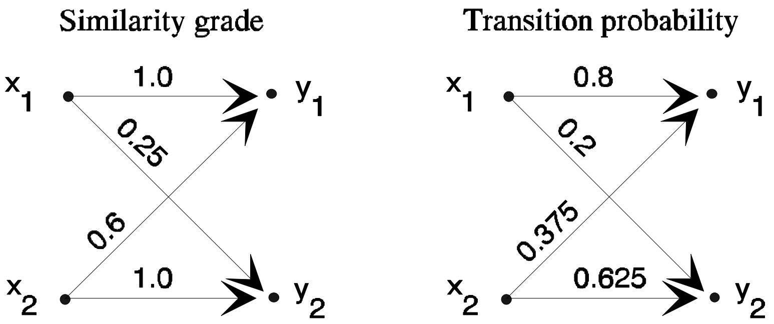


Figure D-4: The Difference between Similarity Grade and Transition Probability (Conditional Probability).

For example, consider Figure D-4 and the two map symbols x_1 and x_2. The grades of similarity $\mu(y_1 \mid x_1) = 1$ and $\mu(y_2 \mid x_1) = 0.25$ are mapped to conditional probabilities as:

$$p(y_1 \mid x_1) = \frac{1}{1 + 0.25} = 0.80 \text{ and } p(y_{21} \mid x_1) = \frac{0.25}{1 + 0.25} = 0.20 .$$

Equation (5) has, in the worst case, the computational effort $T = O(n^2)$ when applied to all similarity classes of X (n is the number of elements of X). Usually, the conflict between map entity x and its neighbouring map entities is limited to the neighbours inside a small region around x. Therefore, we can substitute Y in Equation (5) with

$$Y_{s(x)} = \{ y \in Y \mid y \text{ inside region } s(x) \} ,$$

where $s(x)$ defines a search region around x.

The computation of entropy, equivocation and useful information will be demonstrated from the transition probabilities in Figure D-4. We assume that the probabilities of the two map entities x_1 and x_2 are $p(x_1) = 0.3$ and $p(x_2) = 0.7$. Then the probabilities of the perceived map entities y_1 and y_2 are computed as:

$$p(y_1) = 0.3 \cdot 0.8 + 0.7 \cdot 0.375 = 0.503 \quad \textit{and} \quad p(y_2) = 0.3 \cdot 0.2 + 0.7 \cdot 0.625 = 0.497$$

As a control $p(y_1) + p(y_2) = 1.0$. The entropy of the perceived map entities is computed from Equation (13) (Appendix):

$$H(Y) = -0.503\log_2 0.503 - 0.497\log_2 0.497 = 0.99997$$

The equivocation $H(Y \mid X)$, i.e. the uncertainty in the received signals, is computed from Equation (16) (Appendix). The summations will be broken into small steps which makes it easier to interpret the equation. The uncertainty in the perceived entities when entity x_1 is sent:

$$H(Y \mid x_1) = -0.8\log_2 0.8 - 0.2\log_2 0.2 = 0.72193$$

The uncertainty in the perceived entities when entity x_2 is sent:

$$H(Y \mid x_2) = -0.375\log_2 0.375 - 0.625\log_2 0.625 = 0.95443$$

The mean uncertainty in the perceived map:

$$H(Y \mid X) = 0.3 \cdot 0.72193 + 0.7 \cdot 0.95443 = 0.88468$$

where $X = \{x_1,\ x_2\}$. Finally, the useful information is computed from Equation (21) (Appendix):

$$R = H(Y) - H(Y \mid X) = 0.99997 - 0.88468 = 0.11529$$

Since we have a noisy channel, the entropy of the information source is different from the entropy of the received signals, i.e. $H(X) \neq H(Y)$. The entropy $H(X)$ of the information source is $-0.3\log_2 0.3 - 0.7\log_2 0.7 = 0.88129$ whereas the entropy $H(Y)$ of the received signals is 0.99997.

D.7 ILLUSTRATION OF MAP DESIGN BASED ON INFORMATION THEORY

Bjørke [6] presents a conceptual model for a map design process which incorporates information theory. The model has two main parts: a *map creation* process and a map *evaluation* process. The map creation process creates maps based on knowledge about cartographic design while the map evaluation process evaluates syntactic aspects of the maps based on information theory. The map evaluation process is decomposed into three operational areas. Compared to [6] the three areas will be renamed to:

1) *Source* model;
2) *Stochastic* model; and
3) *Entropy* model.

The source model describes which map information sources, i.e. the map events and their characteristics, are to be selected while the stochastic model describes their stochastic properties as spatial correlation and transition probabilities. Finally, the entropy model uses the source model and the stochastic model to compute different entropy measures as: R, $H(Y)$ and $H(Y \mid X)$. The map design process considered, is presented as the data flow diagram in Figure D-5.

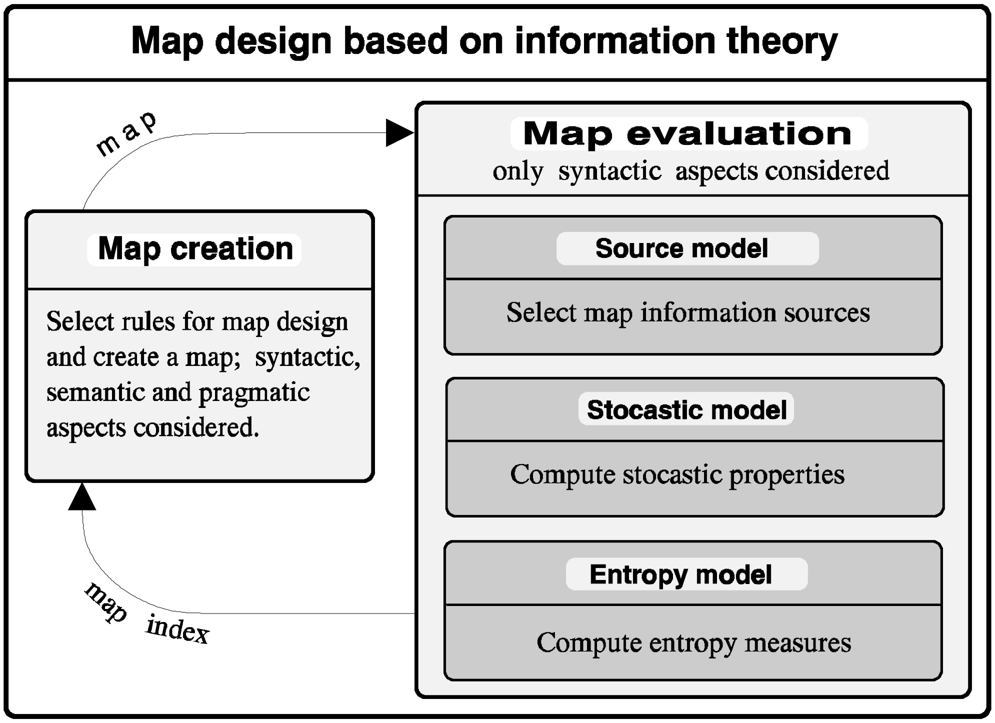


Figure D-5: Map Design Based on Information Theory.

The diagram emphasizes that the map evaluation process considers only syntactic aspects of a map. However, Shannon and Weaver [31] (p. 26) assume that information theory can be applied to all three levels of communication problems. Despite this, in the scope of this paper, the proposed map evaluation process is limited to only syntactic aspects of map information. An automated system based on the proposed map design model is a stepwise procedure. The map creation process generates different maps and thereafter the map evaluation process computes entropy measures for the maps. The information measures (map indexes) are sent to the map creation process, which enables it to draw conclusions about which directions to alter the map design in order to get more efficient maps (the last statement will be elaborated in the examples at the end of this paper). The process cycle of map creation and map evaluation terminates when the map index requirements are met. These requirements should be specified in a sub-process of process map creation (this level of detail is not shown in the figure).

McMaster and Shea [22] describe a map generalization model which decomposes the generalization process into three operational areas:

1) Why to generalize;
2) When to generalize; and
3) How to generalize. Information theory cannot show how to generalize a map, but it is applicable for a better understanding of why and when to generalize or as a tool to control automated cartographic generalization processes.

The three sub-processes of Map evaluation in Figure D-5, mostly cover the operational area (2) of McMaster and Shea.

The relation between the two main processes in Figure D-5 will be elaborated in the context of a map evaluation method described by Morrison [24]. He analyzes the symbolization used on general-purpose atlas reference maps from a semiotics point of view. The simplest definition of semiotics is perhaps "the study of sign systems". In order to systematically evaluate the maps, Morrison [24] states their purpose and concentrates on the semantic and the pragmatic levels of map communication. The application of information theory in the proposed map design process (Figure D-5), can coexist with Morrison's evaluation strategy. Since the Level A evaluations of the information theory method and Morrisons' method evaluate different components of a map, they should not be set against each other. This aspect is considered by the map creation process in Figure D-5, which shows that cartographic knowledge can coexist with information theory in a map design process. With reference to the proposed map design model, Morrison's evaluation method should be applied in the map creation process. Accordingly, information theory evaluations of the syntactic map component together with the map creation process as a whole considers all the three levels of communication problems; syntactic, semantic and pragmatic.

D.7.1 Examples

Some examples will demonstrate the application of information theory to map design. The examples raise several research issues related to map perception. But in order to keep the focus on the application of information theory, detailed discussions of map perception are kept outside the scope of this paper. In the examples a map information source (**Definition 2**) X will be written as:

$$X = (X, C)$$

where X contains the definition of the set of map entities (**Definition 1**) and C represents the characteristic of the map entities. We will use the abbreviations (Top, Met, Pos) for topological, metrical and positional entropies respectively (**Definitions 3**, **4** and **5**). A map information source X for a topological entropy will be written as $X = (X, Top)$ for example. The different entropy measures as $R(X)$, $H(X)$ and $H(Y \mid X)$, which will be used in the examples, are explained and elaborated in the appendix of this paper.

D.7.1.1 Dot Map

Dot maps are often used to show the spatial distribution of discrete geographical point entities. The traditional design rules of dot maps include:

1) Selection of the dot size; and
2) Selection of the number of events per dot.

Figure D-6 shows two dot maps and demonstrates the significance of design rule (2), since map (a) has a lower number of entities per dot than map (b). The evaluation model which will be proposed, will not consider the spatial correlation of the dots. Therefore, a rather simple model can be set up.

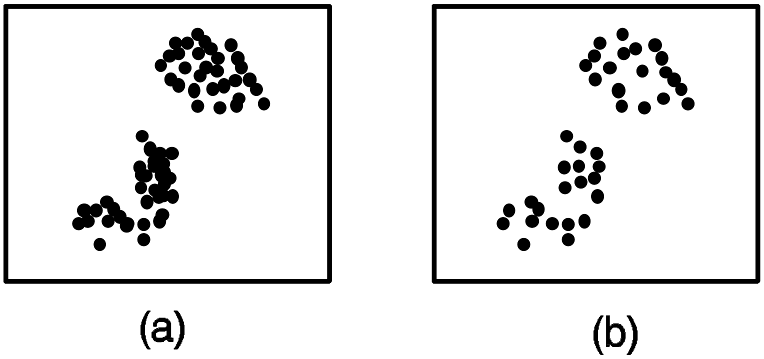


Figure D-6: Two Dot Maps with Different Number of Entities per Dot.

We assume that process map creation (Figure D-5) has set up the following design goal:

- *Design Goal:*
  1) Make the number q of events per dot as small as possible, i.e. as many dots as acceptable to visual perception; and
  2) The preferable dot diameter should be S_0.
- *Source Model:* Select the map information source X for a positional entropy and choose the dots as map entities:

$$X = (X, Pos)$$

$$X = \{x \mid \text{ element } x \text{ is a dot } \}, \text{ i.e. } X \text{ is the set of all dots on the map.}$$

  If aspects of spatial correlation are to be considered in the map evaluation process, the information source for the metrical entropy can be selected as a second information source. In that way we can compute a map index from the two orthogonal information sources (**Definition 6**):

  1) Metrical; and
  2) Positional.

- *Stochastic Model:* The dots are assumed to be equally probable, i.e.:

$$p(x) = \frac{1}{N_x} \text{ for each } x \in X,$$

  where N_x is the number of elements in X.

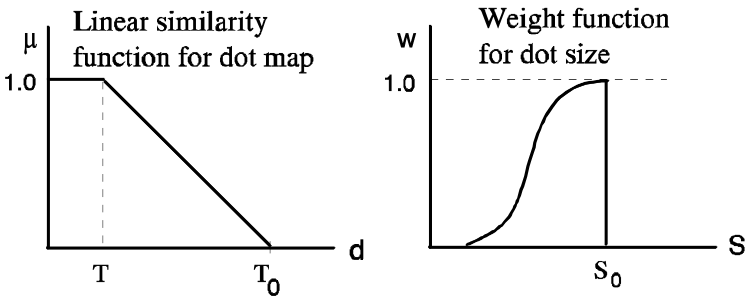


Figure D-7: Functions for Dot Map Design.

The transition probabilities are not as easily derived. Let us assume that we can set up a model so that the visual separation between two neighbouring dots x and y is a function of the distance $d(x, y)$ between them and the dot size S. Further, let us assume that visual separation and visual similarity are inverse quantities. Hence, the similarity function (**Definition 9**) can be defined as $\mu(x, y) = f(S, d(x, y))$. An example of a linear similarity function is given by Figure D-7. In the figure the grade of similarity $\mu(x, y) = 1$, if $d(x, y) \leq T$, i.e. when the dots are so close to each other that they cannot be separated. If $d(x, y) \geq T_0$, the dots are clearly separable and the grade of similarity $\mu(x, y) = 0$. When the similarity function is defined, the transition probabilities which we need for the entropy computations, can be derived from Equation (5). The design of the similarity function should consider the resolution and the type of the output media, the colour of the dots and other parameters related to map perception. A more detailed discussion of this specific topic lies outside the scope of this paper.

- *Entropy Model:* We assume a map creation process $M(q, S)$ which produces dot maps by varying the number q of events per dot and varying the dot size S. Statement (1) of the design goal can be modelled by $\max[R(X) \mid M(q, S)]$, which is the maximum value of the useful information $R(X)$ of map information source X under the constraint that the different map alternatives are produced by $M(q, S)$. Statement (2) of the design goal can be satisfied by a weighted entropy computation. Hence, the map index K can be computed from $K = w(S) \cdot R(X)$; where $w(S)$ is a weight function which takes the dot size as variable. An example of a weight function is given in Figure D-7. From the figure, we can see that the weight has its maximum value at the preferred dot size S_0, i.e. $w(S_0) = 1$. There is no good reason to consider a dot size greater than S_0. Therefore, the weight function is designed so that $w(S) = 0$ when $S > S_0$.

- *Selection Criterion:* Select the pair (q, S) which corresponds to the maximum value of the map index, i.e.:

$$K_{max} = \max[w(S) \cdot R(X \mid M(q, S))],$$

where $R(X)$ is computed from Equation (21) (Appendix). In order to reduce the computational effort of the process cycle, some strategy to eliminate maps which are not candidates to the best solution, should be implemented.

D.7.1.2 Contour Map

An information theory approach to the selection of an appropriate contour interval in contour maps is proposed in [10]. The following example will elaborate this proposal.

- *Design Goal:* Make the contour interval e as small as possible, i.e. as many contour lines as acceptable to visual perception.
- *Source Model:* Select the map information source X for a positional entropy and choose the contour lines as map entities:

$$X = (X, Pos)$$

where X is the set of contour lines of the map.

- *Stochastic Model:* It seems reasonable that a long contour line should have a higher probability than a short contour line. Therefore, we will select the model:

$$p(x) = \frac{l(x)}{\sum_{x \in X} l(x)} \quad \textit{for each } x \in X$$

where $l(x)$ is the length of contour line x and $\sum$ computes the total length of all the contour lines. We assume that the similarity between neighbouring contour lines can be modelled by a function of the type used in the dot map example, i.e. the similarity between two lines is zero when the distance between the lines is greater than T_0. With the exception of parallel lines, the distance between two contour lines will vary. Therefore, the similarity between two contour lines can be computed as a mean value for different sections of the lines.

- *Entropy Model:* We select a map process $M(e)$ which creates maps with contour interval e under the constraint $e \in E$. The constraint can for example limit e to values which are easy to remember. The design goal "as many contour lines as acceptable to visual perception", will be evaluated against the useful information of the maps, i.e. the map index K is computed as $K = R(X)$.
- *Selection Criterion:*

$$\textit{Select the contour interval which corresponds to} \quad K_{max} = \max[R(X \mid M(e))]$$

An experiment based on the model above, was carried out on a digital terrain model of a small part of Norway. Table D-1 summarizes the experiment and shows the map index at different contour intervals for some selected maps.

The computations in Table D-1 assume the map scale 1:120 000; and a linear similarity function (Figure D-7) with $T = 0.1$ mm and $T_0 = 0.4$ mm. In this experiment the map index reached the maximum value 3.275 at contour interval 49 m.

Table D-1: Map Index at Different Contour Intervals – Map Scale 1:120 000.

Contour Interval	Map Index	Entropy	Equivocation
m	K = R(X)	H(Y)	$H(Y \mid X)$
150	2.254	2.254	0.000
125	2.519	2.520	0.001
100	2.833	2.854	0.021
75	3.150	3.292	0.142
60	3.269	3.628	0.359
55	3.273	3.757	0.484
49	**3.275**	**3.918**	**0.643**
48	3.260	3.951	0.691
46	3.244	4.013	0.769

One should note that the selection of parameter values in the similarity function has great influence on the equivocation computation, i.e. the computation of $H(Y \mid X)$. A more detailed discussion of this issue related to map perception, is outside the scope of this paper.

The experiment demonstrates a property of the evaluation method. At a high value of e the contour lines can easily be separated, but there are few of them. On the other hand at a low value of e we have the opposite situation. Our model considers this property of the maps and makes a balanced selection between grade of entropy and grade of equivocation. At the optimum value of e, we have in Table D-1 the equivocation 0.643 and the entropy 3.918; which corresponds to the maximum value of $K = 3.918 - 0.643 = 3.275$. Hence, a property of our selection criteria is that the optimum choice is not necessarily a map with zero equivocation.

D.7.1.3 Line Generalization

An information theory approach to the selection of appropriate parameters in line generalization algorithms, is presented in [5]. The approach presented selects a source model based on angular change. Saga [29] discusses this approach and shows that it is too simplistic to base the selection of generalization parameters on angular change only. Structural information should be considered as well. The complexity of line generalization and the fact that a number of fundamental problems are still unsolved, are pointed out by several authors [21][34].

The present example, deals with line simplification. “Simplification is necessary to eliminate unwanted details (such as small wobbles along lines) that would be difficult or impossible to perceive after scale reduction” [34]. The problem we will put into focus, is how to set up a map evaluation model that can assist us in the selection of an appropriate grade of simplification, i.e. the parameter values of the simplification algorithm. Our example will not be connected to a specific line simplification algorithm, since in principle any simplification algorithm can serve as a basis for the map creation process. The model, which soon will be presented, is not complete since it still is under investigation. Hopefully, it can be looked upon as an innovative framework for further research in this are. We assume the following design goal:

- *Design Goal:* Keep as much of the variation along the line as is acceptable to visual perception.

- *Source Model:* How to compute entropy measures for a line, is not a trivial task. However, we will select two information sources. The first, A is an information source for a metrical entropy; which takes the break angles of the digitized line as map entities. The second, X is an information source for a positional entropy.

$$A = (A, Met)$$

$$X = (X, Pos)$$

$$\begin{aligned} A &= \{\alpha \in R \mid element\ \alpha\ is\ a\ break\ angle\ of\ the\ digitized\ line\} \\ X &= \{x \mid element\ x\ is\ a\ \delta - circle\ of\ the\ line\} \end{aligned}$$

The elements of X are some derived entities, which we term δ-circles (Figure D-8). The concept of $\delta -$circle is demonstrated in Figure D-8. The circles are of equal size and distributed along the line according to the following rules:

1) The circle centers are located on the line;
2) The distance d between the circle centers is constant when measured along the line; and
3) The diameter δ of the circles is equal to d.

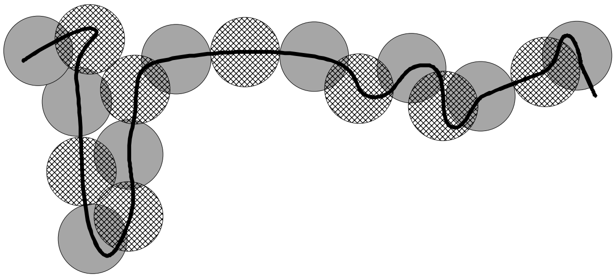


Figure D-8: δ-Circles of a Line.

The diameter of the $\delta -$circles will influence the values of the entropy measures of X. An appropriate circle size is supposed to consider visual limitations as least perceivable winding, etc. However, this specific topic related to the size of δ-circles is an issue for further research.

- *Stochastic Model:*

$$p_\alpha = (p(\alpha) \mid 0 \le \alpha \le 2\pi)$$

where p_α is a probability distribution.

$$p(x) = \frac{1}{N_x} \quad for\ each\ x \in X$$

where N_x is the number of δ-circles of the line. The transition probabilities of the two information sources can be computed using a strategy similar to that of the dot map example.

- *Entropy Model:* We select a map process $M(t)$ which creates different versions of lines by varying the generalization parameter t. The map index is computed as a weighted sum for the R-values of the two map information sources:

$$R = w_\alpha \cdot R(A) + w_x \cdot R(X)$$

where the w's are some weights. Since α is a continuous varying variable, the entropy computation is based on:

$$H(\alpha) = -\int_0^{2\pi} p(\alpha)\log_2 p(\alpha)d\alpha \approx -\sum_{i=1}^{n} p(A_i)\log_2 p(A_i)$$

where the approximation of the integral is based on dividing the continuous domain in n discrete classes, i.e. $A = \cup_{i=1}^{n} A_i$.

- *Selection Criterion:* We assume that the design goal is met at the maximum value of the map index:

$$K_{max} = \max[w_\alpha \cdot R(A \mid M(t)) + w_x \cdot R(X \mid M(t))]$$

The scope of the present model is not to give a complete set of constraints to control the complex line generalization process, but rather demonstrate properties of information theory. Therefore, the information theory model presented, calls for further research in order to achieve a successful cartographic adaptation.

D.7.1.4 Choropleth Map

A statistical surface can be visualised in several ways. One such method is a choropleth representation [27]. The traditional design rules of choropleth maps include:

1) Selection of the number of classes; and
2) Determination of class limits.

In [5] is presented an information theoretic approach to compute an optimum number of classes in choropleth maps. Based on this proposal, the following model is set up:

- *Design Goal:* Select as many classes as acceptable to visual perception, i.e. seek an optimal solution for how much variation of the statistical surface that can be portrayed on the map.

- *Source Model:* Assume a raster map. Select the map information source X for a topological entropy and use the edge between two neighbouring pixels as the map entity:

$$X = (X, Top)$$

$$X = \{x_{lr} \mid (l,r) \in H^2 \wedge (element\ x\ is\ an\ edge\ between\ two\ adjacent\ \ pixels)\}$$

where x_{lr} represents a pair that has the two components: the colour of the left hand pixel and the right hand pixel of x, H is the set of different colours of the map and H^2 is the Cartesian product $H \times H$. If the map has black and white pixels only, we have: $H = \{black, white\} = \{b, w\}$. The set of entities in this case: $X = \{x_{bw}, x_{bb}, x_{wb}, x_{ww}\}$.

- *Stochastic Model:*

$$p(x_{lr}) = \frac{N(l,r)}{\sum\limits_{l \times r \in C^2} N(l,r)} \quad \textit{for each } (l,r) \in H^2$$

where $N(l,r)$ is the number of edges with the colour attribute (l,r).

- *Entropy Model:* Consider the map process $M(h)$ which generates choropleth maps with different number h of classes. The map index to be computed is: $K = R(X \mid M(h))$.
- *Selection Criterion:*

$$\textit{Select the number of classes which corresponds to} \quad K_{max} = \max[R(X \mid M(h))].$$

In [5] the transition probabilities were estimated from an investigation in which thirty subjects were asked to distinguish between different grey values on some test plates. Based on the transition probabilities from the investigation above, entropy measures are computed for two choropleth maps; one map has a correlated spatial distribution of the classes, while the other map has a random spatial distribution of its classes. The entropy, equivocation and the useful information are computed at different class numbers. Table D-2 shows the results of the computation. The table shows that the correlated map gets its maximum value of $R = 2.34$ in 5 classes while the random map gets its maximum value of $R = 2.61$ in 4 classes.

Table D-2: Map Statistics at Different Class Numbers
(The bold face numbers indicate the level of the channel capacity)

The Correlated Map			
Class No.	$H(Y)$	$H(Y \mid X)$	R
3	2.04	0.16	1.88
4	2.71	0.51	2.20
5	**3.32**	**0.98**	**2.34**
6	3.92	1.71	2.21

Table D-2: Map Statistics at Different Class Numbers (cont'd)
(The bold face numbers indicate the level of the channel capacity)

The Random Map			
Class No.	$H(Y)$	$H(Y\mid X)$	R
3	2.48	0.16	2.32
4	**3.24**	**0.63**	**2.61**
5	3.84	1.28	2.56
6	4.33	2.29	2.05

D.7.1.5 Area Elimination

Elimination routines can be used to simplify area features. The criteria may be:

1) Minimum feature size; or
2) Proximity to neighbouring features [27].

Assume a map with equally sized area features (Figure D-9). Due to exaggeration, as a part of map generalization, the map symbols may overlap or may be very close to each other. This is often a problem in small scale maps, which is the case for house symbols in the 1:50 000 topographic maps from the Norwegian Mapping Authority.

- *Design Goal:*
    1) Keep as much of the variation as acceptable to visual perception; and
    2) Eliminate features by proximity to neighbouring features.

- *Source Model:* Select the map information source X for a positional entropy and choose the area features as map entities:

$$X = (X, Pos)$$

$$X = \{x \mid elementxisanareafeature\}$$

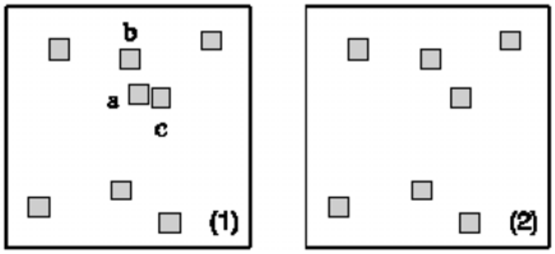


Figure D-9: Simplification by Area Elimination.
(Feature a is a candidate for elimination in map (1) – This feature is eliminated in map (2))

- *Stochastic Model:* Since the features are assumed to be of equal size, their probabilities are modelled as:

$$p(x) = \frac{1}{N_x} \mathrm{mod}\, 5 mmforeachx \in X$$

where N_x is the number of features. The transition probabilities can be derived similarly as in the dot map example.

- *Entropy Model:* Requirement (2) of the design goal can be met by eliminating the feature ε which has the greatest local equivocation, i.e. ε corresponds to $\max_{x \in X}[H(Y \mid x)]$. Requirement (1) can be met by maximizing $R(X)$. Therefore, two map indexes should be sent to the map creation process, one index which corresponds to the local equivocation $H(Y \mid \varepsilon)$ and another index corresponding to the useful information $R(X)$ of the map as a whole.

- *Selection Criterion:*

$$\textit{eliminate the feature } \varepsilon \textit{ which corresponds to } K_\varepsilon = \max_{x \in X}[H(Y \mid x)]$$

$$\textit{select the map which corresponds to } K_{max} = R(X \mid M(\varepsilon))]$$

where $M(\varepsilon)$ is a process which eliminates from the map the feature ε. The computation of $H(Y \mid x)$ is based on the last $\sum$ in Equation (16) (Appendix): $H(Y \mid x) = \sum_{y \in Y} p(y \mid x) \log_2 p(y \mid x)$. For each time an area feature is eliminated from the map, a new candidate to be eliminated should be computed. The process terminates when $R(X)$ receives its maximum value.

The computation of a candidate to be eliminated, will be illustrated. Related to Figure D-9, assume the following similarities: $\mu(a \mid a) = \mu(b \mid b) = \mu(c \mid c) = 1$, $\mu(a \mid b) = \mu(b \mid a) = 0.1$ and $\mu(a \mid c) = \mu(c \mid a) = 0.4$. All other similarities are assumed to be zero. The corresponding transition probabilities are computed from Equation (5): $p(a \mid a) = 0.667$, $p(b \mid a) = 0.067$, $p(c \mid a) = 0.266$; $p(b \mid b) = 0.909$, $p(a \mid b) = 0.091$; $p(c \mid c) = 0.714$ and $p(a \mid c) = 0.286$. The local equivocations are computed as:

$$H(Y \mid a) = -0.667 \log_2 0.667 - 0.067 \log_2 0.067 - 0.266 \log_2 0.266 = 1.16$$

Similarly, $H(Y \mid b) = 0.44$ and $H(Y \mid c) = 0.86$, which gives the priority list for feature elimination: (a, c, b), i.e. feature a is to be eliminated since it generates a higher local equivocation than c and b.

D.8 APPLICATION OF INFORMATION THEORY IN NATO VISUALISATION RESEARCH

D.8.1 Mathematical Background

The channel capacity C of a map, i.e. the information source, is computed as:

$$C = \max R = \max H(Y) - H(Y \mid X), \tag{6}$$

where R is the useful information of the map, $H(Y)$ is the entropy of the interpreted map and $H(Y \mid X)$ is the amount of confusion, i.e. the equivocation of the received message Y when information source X is used.

The entropy of the interpreted map is computed as:

$$H(Y) = -\sum_{y \in Y} p(y) \log_2 p(y), \tag{7}$$

and the amount of confusion in the interpreted map is derived as:

$$H(Y \mid X) = -\sum_{x \in X} p(x) \sum_{y \in Y} p(y \mid x) \log_2 p(y \mid x). \tag{8}$$

The computation of $H(Y)$ and $H(Y \mid X)$ requires that the transition probability $p(y \mid x)$ is known, i.e. we must know the probabilities that the different map colours are misinterpreted as well as correctly interpreted. Moreover, the probability p(x) must also be known, i.e. the probability that the different symbols occur in the map.

The relation between $p(y)$ and $p(x)$ is derived from:

$$p(y) = \sum_{x \in X} p(x) p(y \mid x). \tag{9}$$

Equation ((8)) can be written as

$$H(Y \mid X) = -\sum_{x \in X} p(x) H(Y \mid x),$$

where

$$H(Y \mid x) = -\sum_{y \in Y} p(y \mid x) \log_2 p(y \mid x). \tag{10}$$

We will term the quantity $H(Y \mid x)$ the local equivocation with respect to map symbol x, i.e. an expression for the equivocation introduced by a single map symbol.

D.8.2 Construction of Hyper Networks Based on the Minimum Entropy Principle

In a research presented at the NATO conference IST-063/RWS-010, Bjørke [8] shows how networks can be generalized to hyper-networks by the application of the minimum entropy principle. A binary matrix representation of the network serves as the starting point for the analysis. The cells in the matrix represent the links between the nodes of the network. If there is a link between any two nodes, the corresponding cell in the matrix is associated the colour white. If there is no link between the nodes, the considered cell is coloured black. It is clear that the order of the rows of the matrix can be changed without altering the meaning of the matrix. In this case the sequence of the rows has impact on the entropy of the matrix. Therefore, interchanging the rows can be used to minimize the entropy. This process can be run for the columns as well. In the reordered matrix nodes that are highly connected will form groups of rows. Grouping the similar rows together means identifying hypernodes of the network. This procedure can be used to transform the original network into a hypernetwork composed of hypernodes and hyperlinks. The procedure can be repeated and in that way the network gradually is transformed to higher and higher level of generalization, see Figure D-10. For further evaluations and study of the method, a software package written in MATLAB is available for the members of the NATO research group as open software.

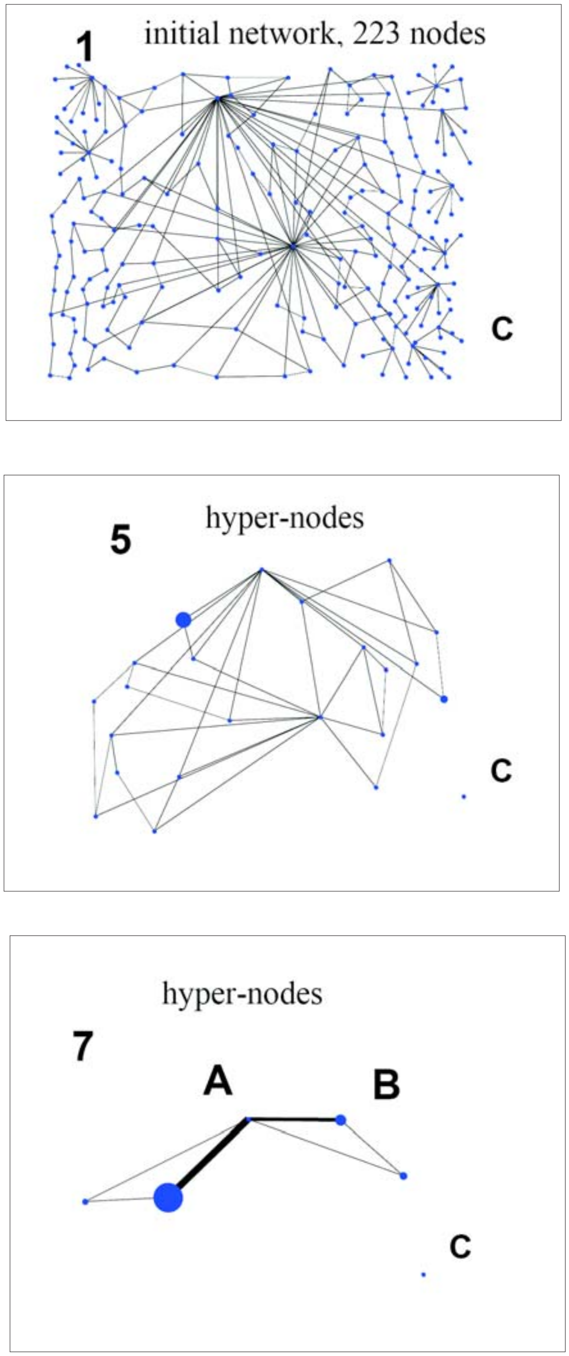


Figure D-10: Illustration of the Information Theoretic Algorithm to Generate Hypernodes.

D.8.3 Constrained Elimination of Links in a Network Visualisation

In a research presented at the NATO conference IST-043/RWS-006 Bjørke [7] shows how information theory can be applied to the generalization of road networks. The idea here is originally presented in [3] and further developed and demonstrated on maps of Norwegian road networks [4]. There are five important elements in the method considered:

1) The computation of the entropy of the network;
2) Computation of the visual conflicts in the network in terms of information theory;
3) The strategy for the elimination of links in the network;
4) Introduction of topological constraints; and
5) The stop criterion based on the computation of the channel capacity of the map.

The elimination of the links in the network is based on Equation (10) by the introduction of the weight function $w(x)$ as:

$$H_w(Y \mid x) = -w(x)H(Y \mid x). \tag{11}$$

Here, $H(Y \mid x)$ represents the local equivocation of each link in the network. The weights are to defined as the inverse of their importance, i.e, higher local equivocation is tolerated from links in important roads than less important roads before they are eliminated. The procedure eliminates the link with the highest value of $H_w(Y \mid x)$ and ends when the stop criterion is reached, see Figure D-11.

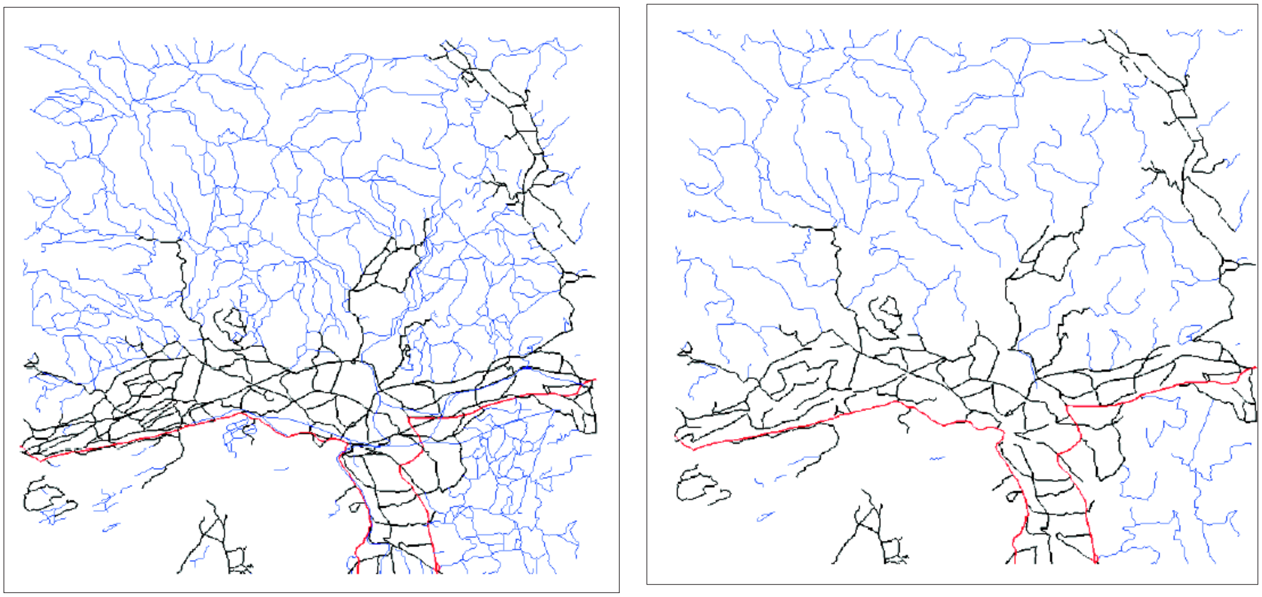

source map
generalized map

Figure D-11: The Information Theoretic Selection Algorithm Demonstrated on Norwegian Road Maps.

D.8.4 Attention Modelling

Research in neorobiology can inspire the development of algorithms that mimic the concept of visual attention. Reynolds [26] concludes that spatial attention causes changes in the neuronal responses that are similar to the effects of increasing the effective contrast of the attended stimulus. This idea can be brought to image design as demonstrated in Figure D-12 and Figure D-13.

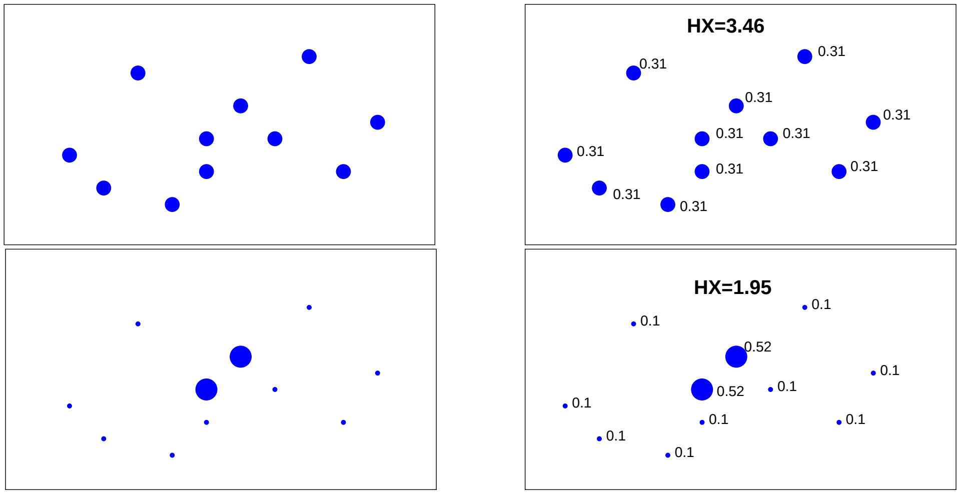


Figure D-12: Mimicking Attention by Use of the Visual Variable Circle Size. The numbers show the information value of each of the symbols. $H(X)$ represents the average entropy of the information source (the image).

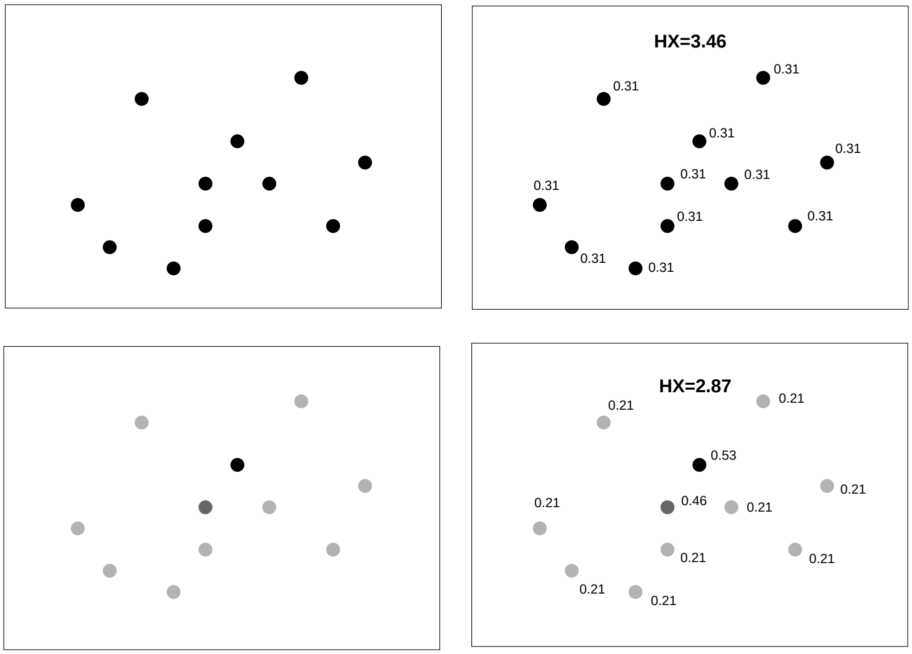


Figure D-13: Mimicking Attention by Use of the Visual Variable Circle Colour. The numbers show the information value of each of the symbols. $H(X)$ represents the average entropy of the information source (the image).

In the first example the circle size is varied and in the next the gray value of the circles are altered. The effect in both cases is that the visual attention is brought to two of the circles. The corresponding entropies are computed and shown in the figures (the average entropy as well as the contribution to this value from each of the map symbols). The question raised here is how to derive the statistical properties of the image. By selecting a rather simple similarity model, the possibilities required for the entropy computation from Equation (7) is derived. The idea here is that a large circle is less similar to it background than a small circle, i.e. when the circle size is small enough the circle cannot be distinguished from its background. Similar for the case with the black and gray circles. In order to map from similarity to probability, a normalization is introduced, i.e. the probabilities must sum to 1.

From Figure D-12 it can be seen that in the case of equal circle size, each of the symbols contributes to the overall entropy with the factor 0.31. In the case of small and large circles the contribution is 0.52 for each of the large and 0.1 for the small circles, i.e. the entropy model is able to catch the effect that the large circles attract the eye to higher degree than the small circles. Therefore, the sketched method is able to model visual attention. A similar result is obtained from Figure D-13. Here, the entropy is computed from the similarity between the background colour and the circle colour. In the case of varying circle colour, the most black circle is assigned the information value 0.53. Then follows the circle with a slightly less black colour with the value 0.46, and finally the other light grey circles with information value 0.21.

A possible utilization of this type of attention modelling is the introduction of entropy thresholds in interactive visualisations. The user can set a entropy threshold and select the most important events. Then the computer system can adjust the symbol size or colour contrast to fit the threshold specified.

A model of the attention process in visual communication is illustrated in Figure D-14. Here, the selector adds priority to the different map symbols, i.e. their relevance. The attention generator changes the visual appearance of the map symbols. The process can be formulated as an optimization. In this way as much of the relevant information is kept in the image and the less relevant information is moved to the back ground color.

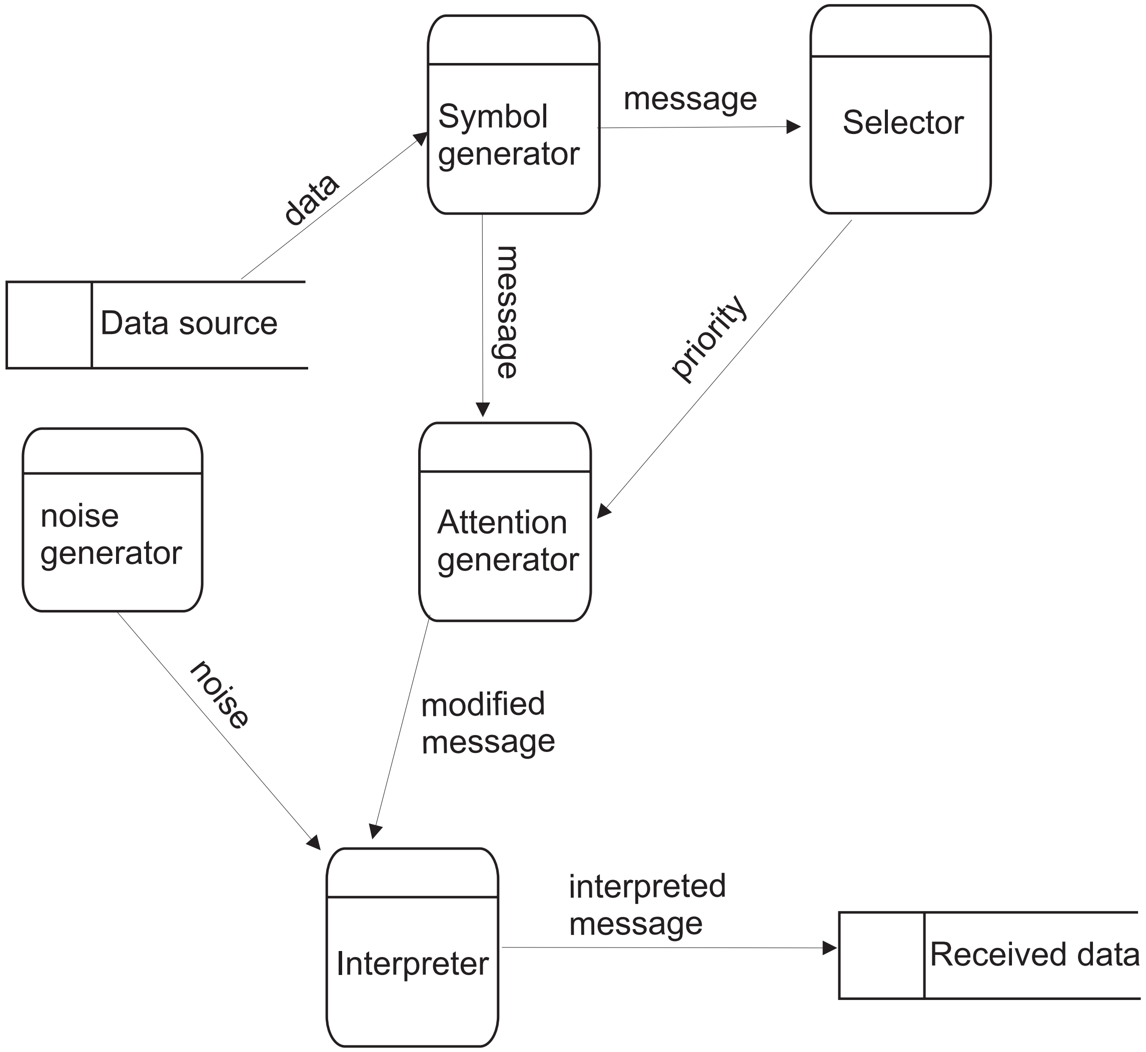


Figure D-14: Attention Modelled as a Data Flow Diagram.

D.9 ACKNOWLEDGMENTS

This research is supported by the FFI-project Poseidon.

D.10 REFERENCES

[1] Bertin, J., “Graphics and Graphic Information Processing”, Walter de Gruyter, Berlin, New York, 1981.

[2] Bjørke, J.T., “Framework for Entropy-Based Map Evaluation”, *Cartography and Geographic Information Systems*, 23(2):78-95, 1996.

[3] Bjørke, J.T., “Generalization of Road Networks for Mobile Map Services: An Information Theoretic Approach”, In *Proceedings International Cartographic Conference (ICA)*, Durban, South Africa, 2003.

[4] Bjørke, J.T. and Isaksen, E., "Map Generalization of Road Networks: Case Study from Norwegian Small Scale Maps", In *Proceedings XXII International Cartographic Conference*, A Coruña, Spain, 11-16 July 2005.

[5] Bjørke, J.T., "Towards a Formal Basis for Cartographic Generalization", In J.T. Bjørke, Editor, *Proceedings Neste generasjon GIS*, pp. 101-109, 7034 Trondheim, Norway, December 1992, The Norwegian Institute of Technology, Department of Surveying and Mapping.

[6] Bjørke, J.T., "Information Theory: The Implications for Automated Map Design", In Milan Konečný, Editor, *Proceedings Europe in Transition: The Context of GIS*, Laboratory on Geoinformatics and Cartography, Department of Geography, Faculty of Science, Masaryk University, Brno, 28-31 August 1994 (The abstract is printed in first edition only).

[7] Bjørke, J.T., "Map Generalization of Road Networks", In *Proceedings of IST-043/RWS-006, Visualisation and the Common Operating Picture*, pp. 1-8, NATO, Research and Technology Organisation, Toronto, Canada, 14-17 September 2004.

[8] Bjørke, J.T., "Reduction of Complexity: An Aspect of Network Visualization", In *Proceedings of IST-063/RWS-010, The Visualising Network Information*, pp. 1-8, NATO, Research and Technology Organisation, Copenhagen, Denmark, 17-20 October 2006.

[9] Bjørke, J.T. and Aasgaard, R., "Cartographic Zoom", In *Proceedings 4th International Symposium on Spatial Data Handling*, Volume 1, pp. 345-353, 23-27 July 1990.

[10] Bjørke, J.T. and Midtbø, T., "Generalization of Digital Surface Models", In *Proceedings 16th International Cartographic Conference*, pp. 363-371, ICA, April 1993.

[11] Bjørke, J.T. and Myklebust, I., "Map Generalization: Information Theoretic Approach to Feature Elimination", In J.T. Bjørke and H. Tveite, Editors, *Proceedings of ScanGIS'2001, 8th Scandinavian Research Conference on Geographical Information Science*, pp. 203-211, Ås, Norway, 25-27 June 2001.

[12] Gatrell, A.G., "Complexity and Redundancy in Binary Maps", *Geographical Analysis*, IX:29-41, 1977.

[13] Gilmartin. P.P., "Influences of Map Context on Circle Perception", *Annals of the Association of American Geographers*, 71(2):253-258, 1981.

[14] Head, C.G., "Mapping as Language or Semiotic System: Review and Comment", In D.M. Mark and A.U. Frank, Editors, *Cognitive and Linguistic Aspects of Space*, pp. 237-262, Kluwer Academic Publishers, Dordrecht, 1991.

[15] Kastner, S., "Attentional Response Modulation in the Human Visual System", In Michael I. Posner, Editor, *Cognitive Neuroscience of Attention*, Chapter 11, pp. 144-156, The Guilford Press, New York, 2004.

[16] Klir, G.J. and Folger, T.A., "Fuzzy Sets, Uncertainty, and Information", Prentice Hall, 1988.

[17] Knöpfli, R., "Informationstheoretische betrachtungen zur verringerung von ungevißheit durch struktur", *International Yearbook of Cartography*, pp. 129-141, 1980.

[18] Knöpfli, R., "Communication Theory and Generalization", In D.R.F. Taylor, Editor, *Graphic Communication and Design in Contemporary Cartography*, pp. 177-218, John Wiley & Sons, 1983.

[19] Lee, D.T. and Schachter, B.J., "Two Algorithms for Constructing a Delaunay Triangulation", *International Journal of Computer and Information Sciences*, 9(3):219-242, 1980.

[20] Li, Z. and Huang, P., "Quantitative Measures for Spatial Information of Maps", *International Journal of Geographical Information Science*, 16(7):699-709, 2002.

[21] Li, Z. and Openshaw, S., "A Natural Principle for the Objective Generalization of Digital Maps", *Cartography and Geographic Information Systems*, 20(1):19-29, 1993.

[22] McMaster, R.B. and Shea, S.K., "Generalization in Digital Cartography", *Association of American Geographers*, 1710 16th Street NW, Washington, DC, 20009-3198, 1992.

[23] Moles, A.A., "Information Theory and Esthetic Perception", University of Illinois Press, Urbana University, 1966.

[24] Morrison, J.L., "Applied Cartographic Communication: Map Symbolization for Atlases", *Cartographica*, 21(1):44-84, 1984, Monograph 31.

[25] Neumann, J., "The Topological Information Content of a Map: An Attempt at a Rehabilitation of Information Theory in Cartography", *Cartographica*, 31(1):26-34, 1994.

[26] Reynolds, J.H., "Attention and Contrast Gain Control", In Michael I. Posner, Editor, *Cognitive Neuroscience of Attention*, Chapter 10, pp. 127-143, The Guilford Press, New York, 2004.

[27] Robinson, A.H., Morrison, J.L., Muehrcke, P.C., Kimerling, A.J. and Guptill, S.C., "Elements of Cartography", John Wiley & Sons, New York, Sixth Edition, 1995.

[28] Robinson, A.H. and Petchenik, B.B., "The Nature of Maps. Essays toward Understanding Maps and Mapping", The University of Chicago Press, Chicago, 1976.

[29] Saga, S.A., "Structural Information in Line Generalization", Unpublished Paper, Department of Surveying, Agricultural University of Norway, 1994.

[30] Salichtchev, K.A., "Some Reflections on the Subject and Method of Cartography after the Sixth International Cartographic Conference", *Canadian Cartographer*, (10):106-111, 1973.

[31] Shannon, C.E. and Weaver, W., "The Mathematical Theory of Communication", The University of Illinois Press, Urbana, 1964.

[32] Simoncelli, E.P., "Vision and the Statistics of the Visual Environment", *Current Opinion in Neurobiology*, 13:144-149, 2003, www.current-opinion.com.

[33] Sukhov, V.I., "Information Capacity of a Map: Entropy", *Geodesy and Aerophotography*, (4):212-215, 1967.

[34] Wang, Z. and Muller, J.C., "Complex Coastline Generalization", *Cartography and Geographic Information Systems*, 20(2):96-106, April 1993.

[35] Weber, W., "Map Generalization – An Information Science Approach", In Harald Opheim, Editor, *Contributions to Map Generalization Proceedings*, pp. 31-52, Oslo, 1980. Norwegian Computing Center. No. 679.

Appendix 1 to Annex D

DA1.1 PROPERTIES OF SHANNON ENTROPY

The principles of Shannon entropy are presented in several textbooks such as [31] or [16]. This section briefly reviews some concepts of Shannon entropy necessary for the development of the theoretical basis of this paper. Given two sets X and Y we can recognize three types of entropies:

- Two *simple entropies* based on marginal probability distribution,

$$H(X)=\sum_{x\in X}p(x)\log_2\frac{1}{p(x)}=-\sum_{x\in X}p(x)\log_2 p(x) \tag{12}$$

$$H(Y)=\sum_{y\in Y}p(y)\log_2\frac{1}{p(y)}=-\sum_{y\in Y}p(y)\log_2 p(y) \tag{13}$$

The maximum entropy is obtained when all events are equally probable, i.e.

$$H(p_1,p_2,\dots,p_n)\le H(\frac{1}{n},\frac{1}{n},\dots,\frac{1}{n})=\log_2 n$$

If the information source is continuous, the entropy computation can be expressed as:

$$H(X)=-\int_{-\infty}^{+\infty}p(x)\log_2 p(x)dx$$

- A *joint entropy* defined in terms of the joint probability distribution on $X\times Y$,

$$H(X,Y)=-\sum_{(x,y)\in X\times Y}p(x,y)\log_2 p(x,y) \tag{14}$$

- Two *conditional entropies* defined in terms of weighted averages of local conditional entropies:

$$H(X\,|\,Y)=-\sum_{y\in Y}p(y)\sum_{x\in X}p(x\,|\,y)\log_2 p(x\,|\,y) \tag{15}$$

$$H(Y\,|\,X)=-\sum_{x\in X}p(x)\sum_{y\in Y}p(y\,|\,x)\log_2 p(y\,|\,x) \tag{16}$$

Based on the relation:

$$p(x,y)=p(y)p(x\,|\,y)=p(x)p(y\,|\,x)$$

it can be shown that:

$$H(X,Y)=H(Y)+H(X\,|\,Y)=H(X)+H(Y\,|\,X) \tag{17}$$

which can be generalized to:

$$\begin{aligned} H(X_1, X_2, X_3, \dots X_n) = \\ & H(X_1) + H(X_2 \mid X_1) + H(X_3 \mid X_1, X_2) \\ & + \cdots + H(X_n \mid X_1, X_2, \dots, X_{n-1},) \end{aligned} \quad (18)$$

It can also be shown that:

$$H(X_1, X_2, \dots, X_n) \le \sum_{i=1}^{n} H(X_i) \quad (19)$$

The equality holds if, and only if, the elements from the n sets are independent in the probabilistic sense. The property of Shannon entropy which follows from Equation (19), is termed the *sub-additive property*. From the rules of probability two sets X and Y are defined as independent if $p(x, y) = p(x) \cdot p(y)$ for each $x \in X$ and each $y \in Y$. If the sets X and Y are independent, their joint entropy is:

$$\begin{aligned} H(X,Y) \quad = \quad & H(\ p(x_1)p(y_1), p(x_1)p(y_2), \dots, p(x_1)p(y_s), \\ & p(x_2)p(y_1), p(x_2)p(y_2), \dots, p(x_2)p(y_s), \dots \\ & \dots, p(x_n)p(y_1), p(x_n)p(y_2), \dots, p(x_n)p(y_s)\ ) \\ = \quad & H(\ p(x_1), p(x_2), \dots p(x_n)\ ) + H(\ p(y_1), p(y_2), \dots p(y_s)\ ) \\ = \quad & H(X) + H(Y) \end{aligned}$$

This property is termed the *additive property* of Shannon entropy.

If the communication channel is noisy, it is not in general possible to reconstruct the original message with certainty by any operation on the received signals. The information loss in a noisy channel is termed *equivocation* and is expressed as a conditional entropy. Let X and Y denote the set of input signals and the set of received signals respectively. The *useful information* R is obtained by subtracting from the source entropy the average rate of conditional entropy (equivocation).

$$\begin{aligned} R \quad &= \quad H(X) - H(X \mid Y) \\ &= \quad H(Y) - H(Y \mid X) \qquad (20)(21)(22) \\ &= \quad H(X) + H(Y) - H(X,Y) \end{aligned}$$

where $H(X \mid Y)$ is the equivocation of the information source when the received signals are known and $H(Y \mid X)$ is the equivocation of the received signals when the signals sent are known. The first expression measures the amount of information sent less the uncertainty of what was sent. The second measures the amount of received information less the part of this which is due to noise. The third is the sum of the entropy of the signals sent and the entropy of the signals received less the joint entropy. The *capacity* C of a noisy channel corresponds to the maximum rate of the transmission and is defined as:

$$C = \max(R) \quad (23)$$

Equation (22) follows when combining Equation (17) and Equation (20) or when combining Equation (17) and Equation (21). The symmetry of Equation (20) and Equation (21) can easily be verified as follows:

Theorem 10 $H(X) - H(X \mid Y) = H(Y) - H(Y \mid X)$

- *Proof:*

$$\begin{aligned}
H(X) - H(X \mid Y) &= H(X) \quad + \sum_{y \in Y} p(y) \sum_{x \in X} p(x \mid y) \log_2 p(x \mid y) \\
&= H(X) \quad + \sum_{y \in Y} p(y) \sum_{x \in X} \frac{p(x)p(y \mid x)}{p(y)} \log_2 \frac{p(x)p(y \mid x)}{p(y)} \\
&= H(X) \quad + \sum_{y \in Y} \sum_{x \in X} p(x)p(y \mid x) \log_2 \frac{p(x)p(y \mid x)}{p(y)} \\
&= H(X) \quad + \sum_{y \in Y} \sum_{x \in X} p(x)p(y \mid x) \log_2 p(x) \\
&\quad + \sum_{y \in Y} \sum_{x \in X} p(x)p(y \mid x) \log_2 p(y \mid x) \\
&\quad - \sum_{y \in Y} \sum_{x \in X} p(x)p(y \mid x) \log_2 p(y) \\
&= H(X) \quad + \sum_{x \in X} p(x) \log_2 p(x) \\
&\quad + \sum_{x \in X} p(x) \sum_{y \in Y} p(y \mid x) \log_2 p(y \mid x) \\
&\quad - \sum_{y \in Y} p(y) \log_2 p(y) \\
&= H(X) \quad - H(X) - H(Y \mid X) + H(Y) \\
&= H(Y) \quad - H(Y \mid X)
\end{aligned}$$

which completes the proof. �

Annex E – NETWORK UNCERTAINTIES

M. Varga, K. Copsey and A. Webb

Uncertainties can be found in the nodes, weights and edges, etc., in networks of any size and characteristic (Refer to Chapters 2 and 3). In this Annex, the application of 'Bayesian Networks' for modeling and reasoning about network uncertainties is discussed. The term 'Bayesian network' is used when referring to probability models and 'network' is used to refer to the communication, social, transportation networks, etc., that are being modeled. The Bayesian network technique can be used to represent and update uncertainties encountered in the network.

E.1 INTRODUCTION

One of the major problems that any decision maker faces is the inherent uncertainty in the data upon which they are trying to base their decisions. Indeed it is believed that the quality of the decision can be improved by understanding uncertainties and the knowledge of how to manage uncertainties. Uncertainties are an un-avoidable accomplice in real problems and there are a wide variety of approaches to handling them; such as fuzzy logic [3], belief functions [6][7], etc. Among these a probabilistic approach has the advantage that it is based upon a formal and rigorous theory.

E.2 PROBABILISTIC MODELS

Probabilistic models based on directed acyclic graphs (DAGs) have a long and established practice. Variants can be found in many different application domains. In cognitive science and artificial intelligence, such models are known as Bayesian networks. They first appeared in the late 1970s due to the need to model the bottom-up and the top-down combination of evidence. The capability for bidirectional inferencing combined with a rigorous probabilistic foundation led to the rapid rise of Bayesian networks as the method of choice for uncertain reasoning in AI and expert systems [1][2][5].

In this appendix a simplified version will be discussed for the purpose of handling uncertainties in networks [9]. The nodes in a Bayesian network represent propositional variables of interest (e.g. location, people, activity) and the links represent informational or causal or non-causal dependencies among the variables, for example:

- The dependence of journey time between two towns on their distance apart.
- Money transaction from one country to another.
- The dependence of weights on height on an individual – clearly not a causal relationship since eating more does not necessarily mean growing taller!

The dependencies are quantified by conditional probabilities for each node given its parents in the network. The Bayesian network supports the computation of the probabilities of any sub-set of variables given evidence or observation about any other sub-set.

In many ways a probabilistic model is a very suitable model for handling network uncertainties as the probabilities provide the natural numerical estimates to weigh the uncertainties and their relationships. More often than not decision makers make decisions based upon the relevant information available and use their previous experience and their intuition to decide upon the uncertainty/probability aspect.

In short, a Bayesian network approach is well suited for analysing networks as it depicts naturally the relationship between various elements, e.g. nodes and links which in turn shows conditional independence and dependence. This provides an important means for deciding the inter-relationship between nodes and links. It is also able to update probability distributions; for example, at a particular instance of a network status and a prior probability distribution over a hypothesis variable that represents a possible network anomaly, the Bayesian network provides the capability to update the probability distribution when there is any change in the network. Furthermore, this technique can be used to analyse the flow of information and uncertainty in the network.

In Chapter 4 a network was briefly described in which there is no link between two nodes, for example in Figure E-1 below the link between nodes 2 and 3 has a value 0 assigned to it and the corresponding cell is coloured white in the matrix. There is no prior knowledge as to why there is not a link between them [8]. Indeed the assumption is that there is nothing at all to preclude there being a link between them.

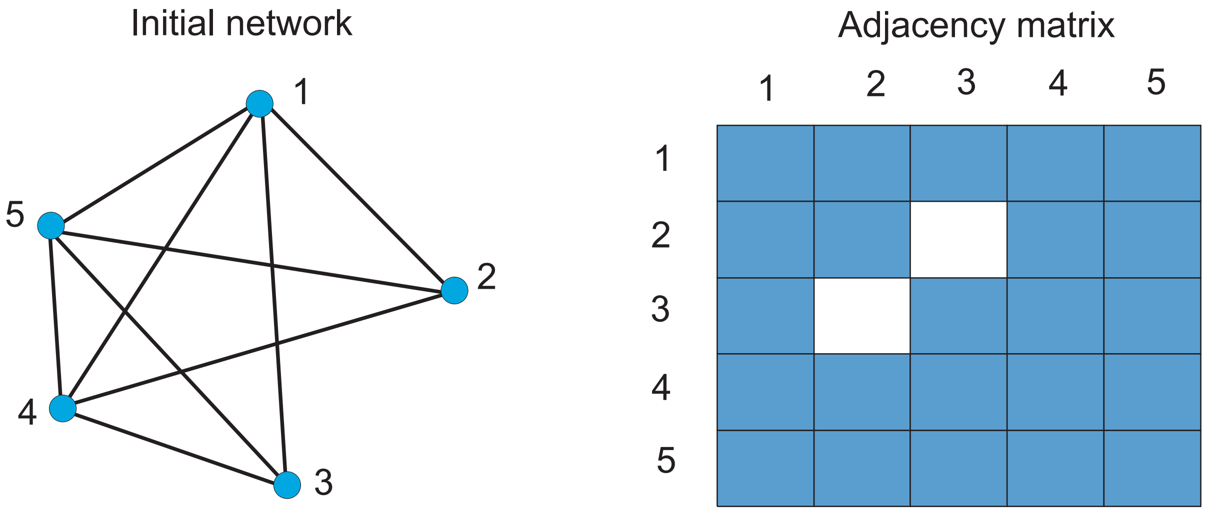


Figure E-1: A Network and its Adjacency Matrix.

There will, however, be cases where some links are prohibited or highly unlikely. The question is how to differentiate between links that just do not exist on the one hand and links that are prohibited on the other hand (because this information will be of importance to the decision maker), and also how to work with this information. There are many different ways that this can be addressed; one way is through the use of prior beliefs (or probabilities), e.g. that a link cannot exist or is highly unlikely, alongside measurements, for example, as in Bayes' theorem. The nodes in the Bayesian network are the edges in the above graph of the network.

Previously it was shown (Chapter 4) that if there is *a priori* knowledge that it is impossible or prohibited (or at least highly unlikely) that nodes 2 and 3 are connected, then it is necessary to differentiate a link of this characteristic from a link that is possible but not observed, such as that between nodes (2,5) and nodes (1,4), see Figure E-2.

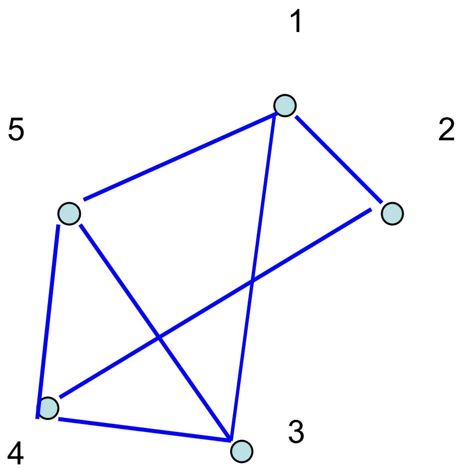


Figure E-2: Network Observations, with No Link between Nodes 2 and 3, Nodes 2 and 5, and Nodes 1 and 4.

Let us define a variable for the relationship between nodes i and j:

- $x_{i,j} = 1$ means that there is a link between nodes i and j;
- $x_{i,j} = 0$ means that there is no link between nodes i and j

and

- $y_{i,j} = 1$ means that a link is observed between nodes i and j;
- $y_{i,j} = 0$ means that a link is not observed between nodes i and j.

We can assign prior probability values for the variable representing the relationship between nodes i and j. For example:

- $p(x_{i,j} = 1) = 0.9$, meaning that we believe (*a priori*) that there will be a link (the closer this probability value is to 1, the greater our prior belief is that there is a link).
- $p(x_{i,j} = 0) = 0.1$, i.e. $p(x_{i,j} = 0) = 1 - p(x_{i,j} = 1)$.

In this example above our prior belief favours a link. In contrast, if $p(x_{i,j} = 0) \approx 1$ then our prior belief is that there is no link.

The following are examples of the observation probability (termed the likelihood) of a link between nodes i and j being observed, which is a function of whether the link actually exists or not.

- $p(y_{i,j} = 0|x_{i,j} = 1) = 0.1$
- $p(y_{i,j} = 1|x_{i,j} = 1) = 0.9$
- $p(y_{i,j} = 0|x_{i,j} = 0) = 0.95$
- $p(y_{i,j} = 1|x_{i,j} = 0) = 0.05$

We can then apply Bayes' Theorem $p(x_{i,j}|y_{i,j}) = p(y_{i,j}|x_{i,j})p(x_{i,j})/\sum_{x_{i,j}} p(y_{i,j}|x_{i,j})p(x_{i,j})$ to update our probability estimates for the relationship variables. For the above example we get:

- $p(x_{i,j} = 1|y_{i,j} = 1) = 0.993$
- $p(x_{i,j} = 1|y_{i,j} = 0) = 0.486$
- $p(x_{i,j} = 0|y_{i,j} = 0) = 0.513$
- $p(x_{i,j} = 0|y_{i,j} = 1) = 0.006$

In contrast, if we set $p(x_{i,j} = 1) = 0.0$ (i.e. our prior probability of a link is zero) then we get:

- $p(x_{i,j} = 1|y_{i,j} = 1) = 0$
- $p(x_{i,j} = 1|y_{i,j} = 0) = 0$
- $p(x_{i,j} = 0|y_{i,j} = 0) = 1$
- $p(x_{i,j} = 0|y_{i,j} = 1) = 1$

In the figure below, link $x_{2,3}$ is prohibited through setting of $p(x_{2,3} = 1) = 0.0$, and the square is coloured black, whereas the square is coloured white when a link $y_{2,5}$ can exist (prior beliefs $p(x_{i,j} = 1) = 0.9$)) but is not observed.

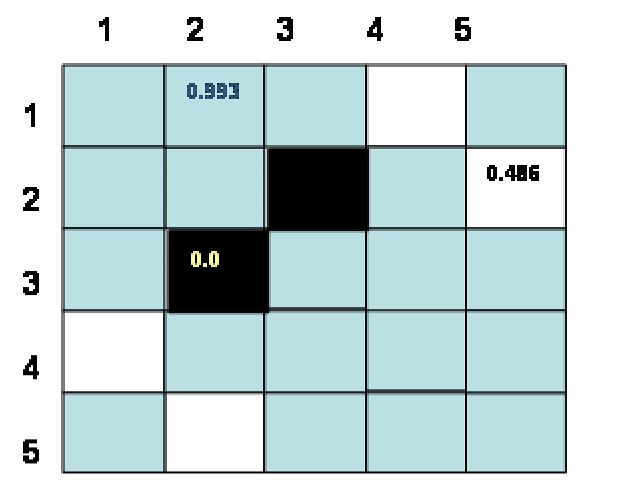


Figure E-3: Posterior Values of Probability of Link Given Observation.

At (2,3) we assume that no link can exist; at (1,4) and (2,5) we believe link would exist though not observed.

We can extend this by applying the prior knowledge of each node in question (if known) so that global behaviour of the whole network can be better represented and understood through the *local* properties of nodes and links. This in turn will benefit the monitoring of network behaviour for anomaly detection and prediction.

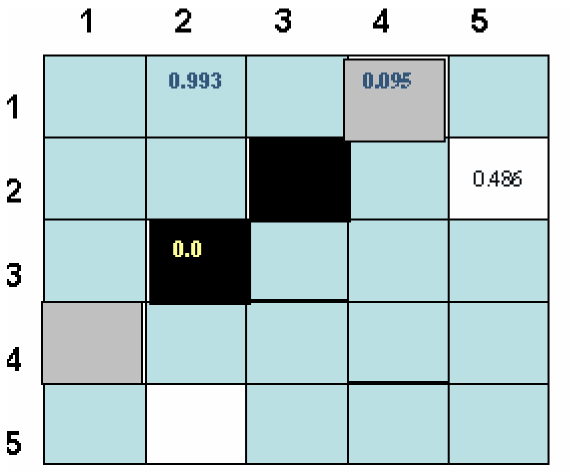
However, in real life we are likely to have limited prior knowledge about all the links and nodes, for example (1,4). In this case we may make the assumption of everything being equally likely *a priori*, i.e.:

- $p(x_{i,j} = 1) = 0.5$
- $p(x_{i,j} = 0) = 0.5$

Then applying Bayes' Theorem $p(x_{i,j}|y_{i,j}) = p(y_{i,j}|x_{i,j})p(x_{i,j})/\sum x_{i,j}p(y_{i,j}|x_{i,j})p(x_{i,j})$ using unchanged observation likelihoods (since only our prior beliefs have changed), to create posterior probabilities for the relationship variables we get:

- $p(x_{i,j} = 1|y_{i,j} = 1) = 0.947$
- $p(x_{i,j} = 1|y_{i,j} = 0) = 0.095$
- $p(x_{i,j} = 0|y_{i,j} = 0) = 0.905$
- $p(x_{i,j} = 0|y_{i,j} = 1) = 0.053$

In Figure E-4, link $x_{1,4}$ has an a prior probability of $p(x_{1,4} = 1) = 0.5$, and the square is coloured grey since the link was not observed. Link $x_{2,3}$ has an a prior probability setting of $p(x_{2,3} = 1) = 0.0$, and all other links have a prior probability of $p(x_{i,j} = 1) = 0.9$. Figure E-5 shows a network representation with uncertainties using the same colour scheme to represent all the unobserved links, i.e. white, grey and black. This is a representation of the 'complete' graph that shows where there are uncertainties and to what degree. This provides an important element for understanding and manipulating network uncertainties and also provides a means to mitigate the uncertainties. This network representation has two advantages, namely the black-grey-white colour scheme provides an intuitively easy means of visualising uncertainties in the network and other colours can be used to represent other aspects of the network properties without the problem of confusion. Other visualisation approaches will be discussed in the next group.



	1	2	3	4	5
1		0.993		0.095	
2					0.486
3		0.0			
4					
5					

Figure E-4: The Posterior Probability Given Observations. It shows links that are possible, probably likely or highly unlikely.

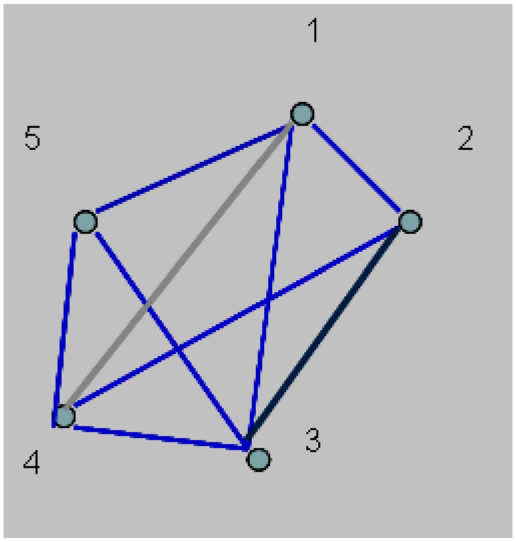


Figure E-5: Network Representation with Unlikely (grey) and Highly Unlikely (black) Links.

The value of $p(x_{i,j} = 1|y_{i,j} = 0)$ for the non-prohibited links is significantly different when the prior knowledge is changed, i.e. it changes from 0.486 to 0.095 as our prior belief decreases.

It is important to note that in this case the observation probability (likelihood) does not change for nodes that are in question. The prior probability ($p(x_{i,j})$), however, does change dependent on which nodes are in question, and therefore the posterior probability $p(x_{i,j}|y_{i,j})$ changes as well. If some node connections are more difficult to observe than others, then the likelihood will also change from node to node.

E.3 PROPAGATION OF UNCERTAINTIES

A Bayesian network is a complete model for the variables and their relationships; it can be used to answer (probabilistic) questions about them. For example, from observations of changes in the network it can be used to identify changes in the state of a sub-set of variables when changes in other node/link variables are observed. This process of computing the *posterior* distribution of variables given observation is called probabilistic inference. The posterior gives a universal sufficient statistic for network behaviour and we can manipulate values for the variable sub-set which reduce or minimize the impact of uncertainties, for instance, the probability of decision error due to insufficient information. A Bayesian network can thus be considered a mechanism for automatically applying Bayes' theorem to complex network problems.

The most common exact inference methods are:

- Variable elimination, which eliminate (by integration or summation) the non-observed non-query variables one-by-one by distributing the sum over the product;
- Clique tree propagation, which caches the computation so that many variables (nodes/links) can be examined/assessed at one time and new observation/evidence can be propagated quickly; and
- Recursive conditioning, which allows for a space-time tradeoff and matches the efficiency of variable elimination when enough space is used.

All of these methods have complexity that is exponential in the network's tree width. The most common approximate inference algorithms are, for example, stochastic MCMC simulation.

E.4 MEASURE OF EFFECTIVENESS (MOE)

The effectiveness of the Bayesian Network is determined by its ability to utilize information to update belief in the observation [2]. How well it performs depends on the functional specification which defines the degree of influence that the information variables have over the observables. Hence, a measure of effectiveness can be obtained, for example, by obtaining a measure of influence, and this can be obtained by a *mutual information* function.

The prior probability distribution P(x) over the observable at any stage of the network monitoring or decision making process reflects the status of network at the time in question. As the situation changes the probability distribution varies. In essence, if the network status is such that P(x) belongs to a sub-set of properties D then the functionally specified part of the network makes maximum use of the information available. P(x) wandering off from D reflects a change (a decrease) in the network's ability to exploit new evidence. To bring back the optimal performance it is therefore necessary to change the functional specification so that with respect to the new network P(x) is back within D. This can be achieved by:

- Assessing the observables so that the information gathered has a better degree of relevance to the situation, i.e. network status.

- Amending the connecting nodes and therefore the links in the network so that the chain of the propagation of the observables is of more direct relevance to the network status in question.

E.5 CONCLUSION

This appendix has provided a brief overview of the use of Bayesian Networks in network modeling. It showed how prior knowledge can be used to model the network behaviour in combination with little or no knowledge from observations and also how, as new data becomes available, it can be used to modify prior beliefs. It can be seen that the Bayesian Network technique is therefore a very powerful tool. The key feature of the use of Bayesian Networks is thus that they enable us to model and reason about uncertainty. This work will be further developed in the new group.

E.6 REFERENCES

[1] Bernado, J.A. and Smith, A.F.M., "Bayesian Theory", John Wiley & Sons, New York, 1994.

[2] Cover, T.M. and Thomas, J.A., "Elements of Information Theory", John Wiley & Sons, New York, 1991.

[3] Cheeseman, P., "Probabilistic vs. Fuzzy Reasoning", In *Uncertainty in Artificial Intelligence*, (Kana, L.N. and Lemmer, J.F., Ed), North-Holland, Amsterdam, 1986.

[4] Fine, T., "Theory of Probability", Academic Press, New York, 1973.

[5] Kanal, L.N. and Lemmer, J.F., "Uncertainty in Artificial Intelligence", North-Holland, Amsterdam, 1988.

[6] Pearl, J., "Fusion, Propagation, and Structuring in Belief Networks", *Artificial Intelligence*, **29**(3):241-288, 1986.

[7] Pearl, J., "Probabilistic Reasoning in Intelligent Systems", *Morgan Kaufmann,* ISBN 0-934613-73-7, 1988.

[8] Varga, M.J. and Bjørke, J.T., "Hyper-Network with Uncertainties", In *Proceedings of the NATO Network-of-Experts Workshop on Network Analysis and Visualization for Simulation and Prediction*, Aerospace Corporation, El Segundo, California, USA, 6-8 November 2007.

[9] Webb, A.R., "Statistical Pattern Recognition", John Wiley & Sons, Chichester, 2nd Edition, August 2002.

Annex F – SURVEY TAXONOMY AND ANALYSIS

J. Treurniet

In this Annex, the details of the technology survey and of the literature search are given. First, the taxonomy used to classify the network visualisation products is described in Section F.1. In this section, statistics are presented for some categories within the taxonomy. To calculate the statistics, the database was loaded into MATLAB and an array of the entries was generated. Statistics were gathered using this array. A large number of fields were empty, indicating that many products had unknown properties or incomplete entries. To complete the entries would be expensive and time-consuming; therefore this rough estimate of the state of the art is accepted. The papers discovered in the literature search are analysed in Section F.2, providing a gross view of the current fields of interest to researchers in information and network visualisation.

F.1 TECHNOLOGY SURVEY TAXONOMY

In an effort to categorize the existing network visualisation products, the taxonomy shown in Figure F-1 was developed. This taxonomy allowed for the development of a searchable database from which gaps could be identified and trends could be observed. There are six first-level categories in the taxonomy: Context, Network Representation, Visual Enhancements, User Interaction, Analysis, and Deployment. The sections to follow describe these categories in more detail.

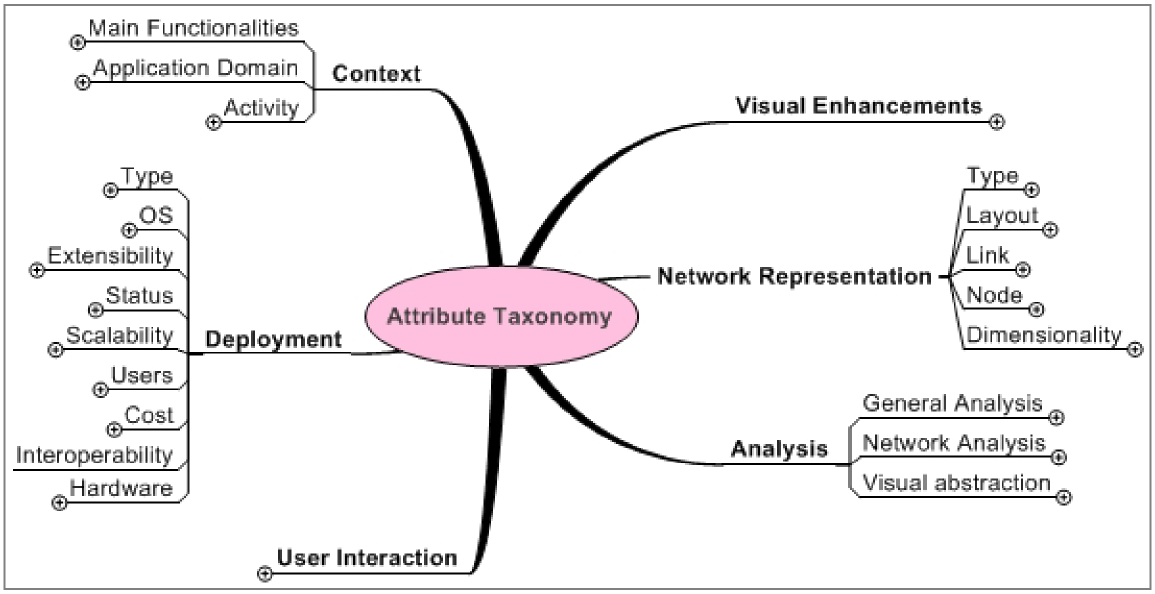


Figure F-1: A Truncated Diagram of the Product Attribute Taxonomy.

It is important to note that due to the scale of the field, an exhaustive product search could not be carried out. As well, commercial products were assessed based only on the available documentation. This led to several empty fields in the survey database. For the full report of the products found in this snapshot in time, the reader is referred to [1].

F.1.1 Context

The first-level category “Context” gives a place where one can describe the general use of the product or tool. Context is divided into three second-level categories: Main Functionalities, Application Domain, and Activity.

The Main Functionalities category is intended to describe the primary capabilities of the software. The Application Domain category is intended to give an indication of the target audience of the software. Activity is divided into three categories: monitor, track and investigate. The taxonomy initially included a “User Role” field, including entries such as “analyst”, “officer” and “chief of staff”, intended to give an idea of the level of detail that the product provides. In the end, this field was left empty for all products and so it was removed from the taxonomy.

In 88% of the cases, the primary functionality of the software was to automate the layout of the network and view it. Fifty-nine percent of the products provided graph manipulation capabilities, and 25% provided network analysis capabilities.

The application domains for the software are as follows, with some products belonging to more than one domain:

- Any domain: 59%
- Computer networks: 27%
- Social networks: 16%
- Biology: 2%
- Databases: 1%

F.1.2 Network Representation

The first-level category “Network Representation” is used to describe how the software displays the information visually.

F.1.2.1 Type

The second-level category “Type” indicates the type(s) of network that the software supports, e.g. acyclic, stigmergic, hierarchical or non-hierarchical, directed or undirected, planar or non-planar, multimodal.

Only 23 of the product entries specified the type of graph that the product supports. Of these, 14 specified both directed and undirected, 5 handled only directed, and 2 handled only undirected. Only one product specified that it could display multimode graphs, and one product listed semantic networks.

F.1.2.2 Links

The “Links” category identifies the attributes for network links that can be modified by the user. These include the attributes: traffic, weight, label, user-defined, pre-defined attributes, and colour.

F.1.2.3 Nodes

The “Nodes” category identifies the attributes for network nodes that can be modified by the user. These include: symbol, label, nested, user-defined, pre-defined attributes, and colour.

F.1.2.4 Layout Algorithms

Although the layout algorithms are all included in the database as direct children of the "Layout Algorithms" category, several of the algorithms are quite similar. Rather than list each individual layout by name, the layouts are categorised further in this section, into 12 layout algorithm types: simple layouts, algorithms for rooted trees, algorithms for general directed graphs, algorithms for free trees, algorithms for planar graphs, force directed algorithms, circular layout algorithms, grouped by attribute similarities, machine-learning algorithms, extensions, information display, and other techniques.

Figure F-2 shows a histogram of the layout algorithms. It appears from this histogram that the only areas under-represented in the products are layouts for free trees and machine-learning algorithms. However, force-directed methods may be applied to free trees.

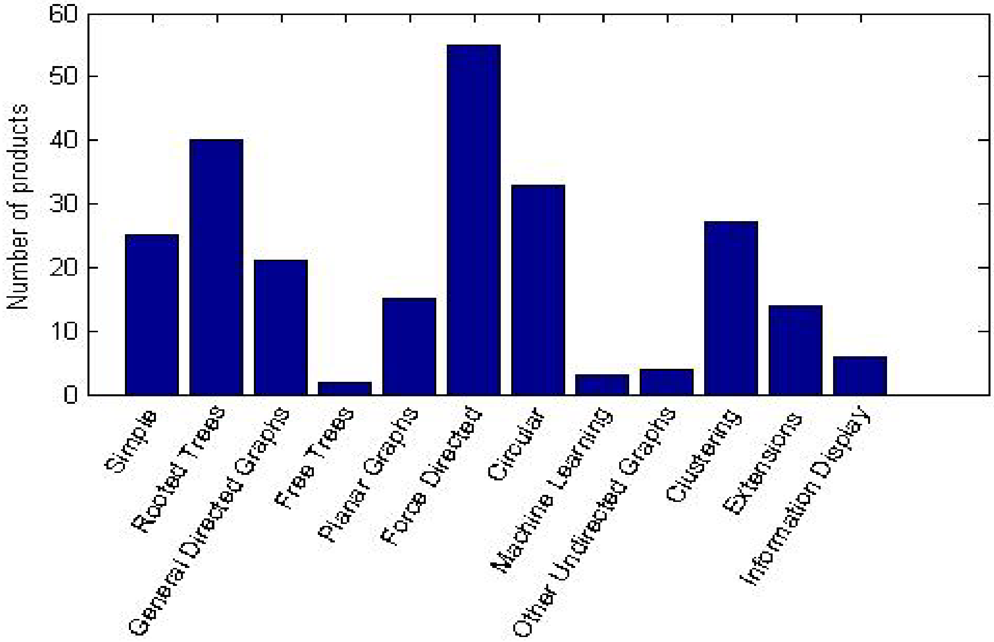


Figure F-2: The Distribution of Network Layout Capabilities in the 139 Products in the Survey.

In the following sections, the layouts named in the product survey are listed according to the 12 categories.

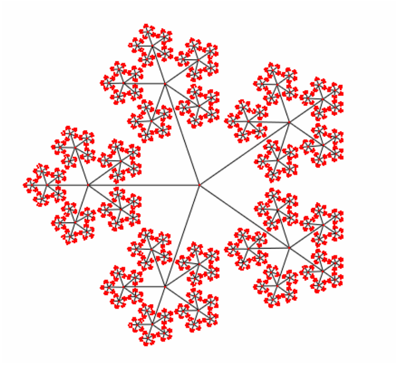
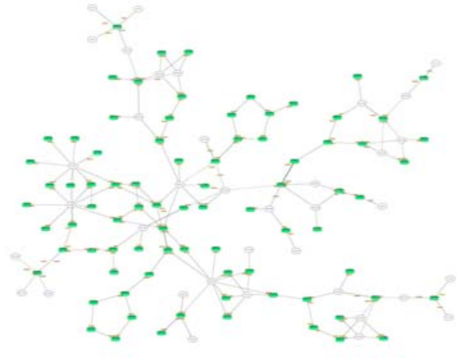
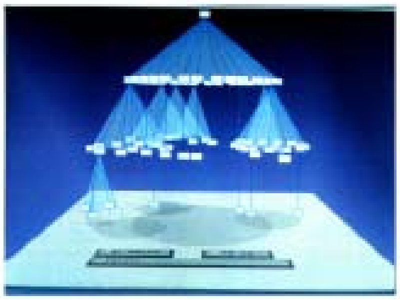
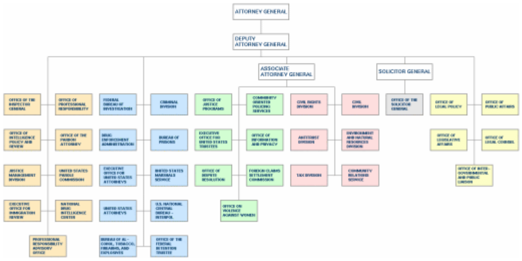
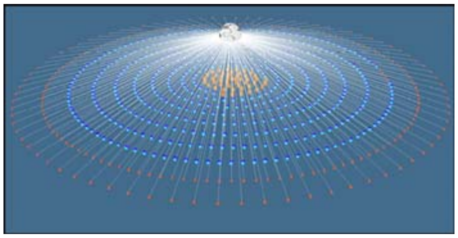
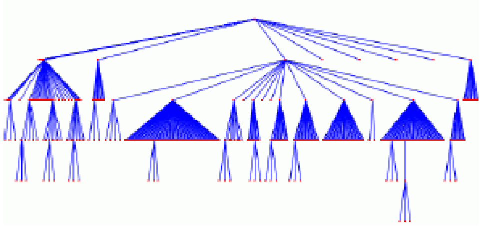
F.1.2.4.1 Simple Layouts

Simple layouts are layouts that require little or no computation, and may not show links. These include coordinate-based placement, grid-based placement, and random placement.

F.1.2.4.2 Algorithms for Rooted Trees

A *rooted tree* is a hierarchical graph that has no crossings and no loops. A selection of examples for this category is shown in Table F-1.

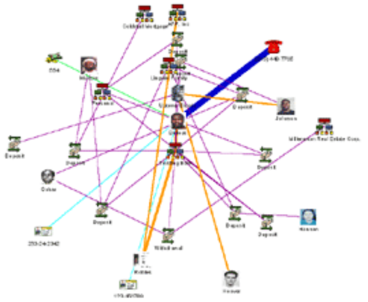
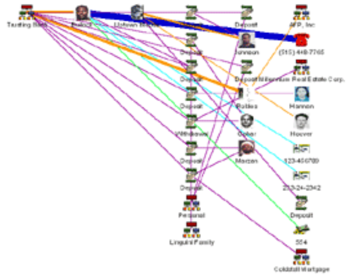
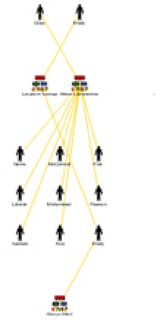
Table F-1: A Selection of Examples of Algorithms for Rooted Trees

 Bubble Tree [2]	 Symmetric [3]
 Cone Tree [4]	 OrgChart [5]
 Nested Cone Tree [6]	 Reingold-Tilford [7]

F.1.2.4.3 Algorithms for General Directed Graphs

A *general directed graph* is hierarchical and may have crossings or loops. Sugiyama's algorithm falls into this category, as well as Centrality Placement and Level Span (see Table F-2).

Table F-2: General Directed Graph Layout Examples

 Centrality Placement [8]	 Centrality Placement [8]	 Level Span [8]

F.1.2.5 Algorithms for Free Trees

Free trees are non-hierarchical graphs that have no crossings and no loops. Although force-directed algorithms may be applied to these graphs, they are awarded a section of their own due to their applicability to general undirected graphs. This category includes the Tutte barycentre placement algorithm: starting from an order on the top and bottom layers, the coordinates of a node are defined to be the barycenter of those of its neighbours. This corresponds to the intuitive idea that a node should be kept "close" to its neighbours. The solution is then obtained by solving a system of linear equations [7].

F.1.2.5.1 Algorithms for Planar Graphs

A *planar* graph is a graph that is non-hierarchical, has no edge crossings, and may have loops. Layout algorithms include Bus, Mixed Model, Orthogonal, FPP, Schnyder, and Ring layouts, among others.

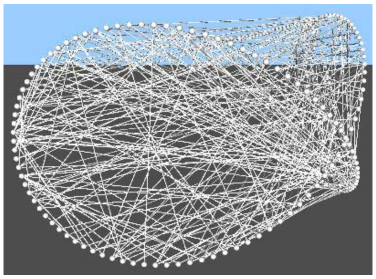
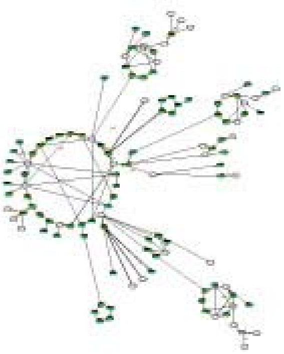
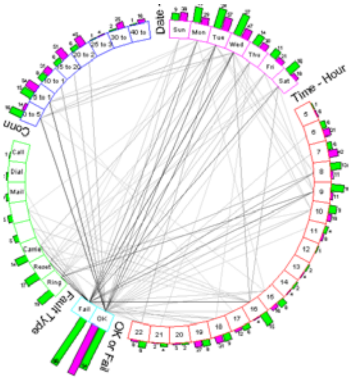
F.1.2.5.2 Force Directed Algorithms

Force Directed methods may be applied to free trees or general undirected graphs (which are non-planar and may or may not be cyclic). The common thread among these algorithms is that they are based on physical models, and are placed where the total energy is minimized within the system. Some examples of force-directed algorithms are: Eades' Spring algorithm [9], the Fructerman-Reingold Spring algorithm [10], the LinLog layout [11], the Kamada-Kawai layout [12], GEM [13], and the ACE [14] and GRIP [15] algorithms for large graphs.

F.1.2.5.3 Circular Layout Algorithms

Circular layouts can be applied to any non-hierarchical graph data set (i.e. a general undirected graph). Examples of circular layouts are shown in Table F-3.

Table F-3: A Selection of Circular Layouts

 Butterfly [16]	 Circular [3]	 Daisy Chart [17]

F.1.2.5.4 Machine Learning Algorithms

These algorithms use machine learning techniques to determine the optimal placement of nodes. Only the Inverted Self-Organizing Map [18] was discovered in the product survey.

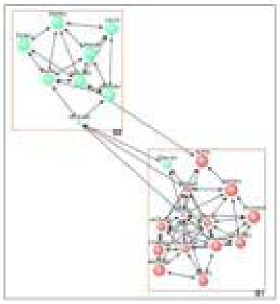
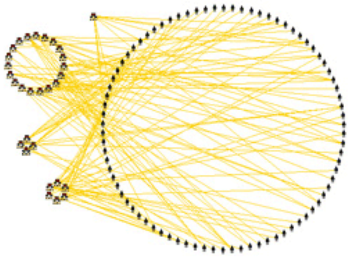
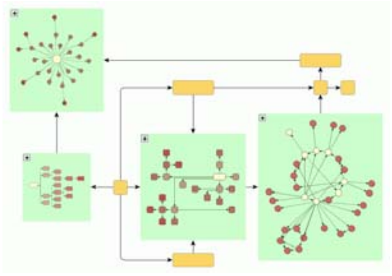
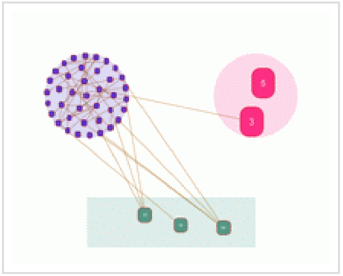
F.1.2.5.5 Other Techniques for Undirected Graphs

This section is reserved for the layout algorithms that do not have a home in the preceding sections. This includes high-dimensional embedding (e.g. [19]), spectral graph drawing (e.g. [20]), and the topological graph layout [21].

F.1.2.5.6 Grouped by Attribute Similarities

There are several products capable of placing nodes with similar attributes together, with different names for similar layouts. Some examples of these layouts are shown in Table F-4.

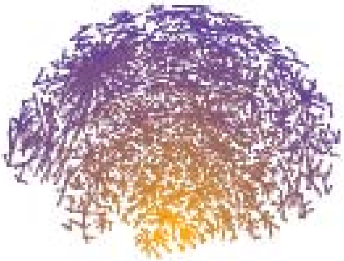
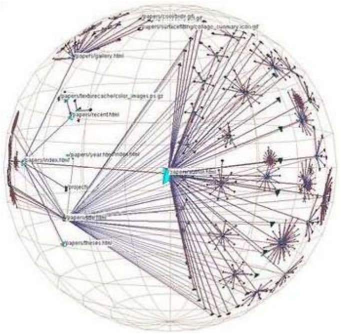
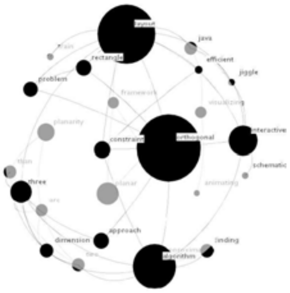
Table F-4: Examples of Clustered Layouts

 Orthogonal Cluster Layout [3]	 Cluster Analysis [24]	 Group By [8]
 Nested Graph Hierarchies [23]	 Weighted [8]	 Zoned [22]

F.1.2.5.7 Extensions

Extensions are algorithms that are applied to the layout after the nodes and links have been laid out in some initial fashion. Examples of extensions are shown in Table F-5. Additionally, this category includes the classification term "incremental", which means that very little changes in the node placements when a new node is added or a node is deleted. Four products included this capability [21][23][25][26].

Table F-5: Examples of Extension Algorithms

 Topological Fisheye [27]	 3-D Hyperbolic View [28]	 Spherical [29]

F.1.2.5.8 Information Display

These are not network layouts, but provide alternate views or a means of viewing information about a network that is not conveyed through the nodes and links. Information displays found in the products were parallel coordinates, treemaps, tabular presentation of data, and time-evolution plots.

F.1.2.6 Dimensionality

The geometric dimensions that the software supports for analyzing/display networks may be one of: 2-D, 3-D, geospatial, temporal. Of 120 products with entries for the dimensionality field, 89% of the products used 2 dimensions, 30% of the products used 3 dimensions, 16% of the products were capable of geospatial displays, and 13% of the products provided temporal displays.

F.1.3 Analysis

The Analysis section is divided into "Network Analysis" (a sub-set of graph theory), "Visual Data Abstractions", and "General Analysis".

- **Network Analysis:** The methods that the software provides for analyzing the properties of networks (or graphs). These properties are generally domain independent and are used to describe qualities of the graph that may be useful in further analysis as it applies to a specific domain; however some metrics are currently only applied in fields such as social networks or computer networks. Definitions of network analysis terms can be found in Annex I.
- **Visual Data Abstraction:** Specifies any additional visual methods used for further analyzing data (e.g. charts and scatter plots).
- **General Analysis:** Additional analysis methods and techniques that do not fit into the "Network Analysis" category such as statistics/metrics and data transformation.

F.1.3.1 Network Analysis

For the statistical analysis of this section, the terms used in the database were catergorised according to four types of analysis: centrality measures, cluster recognition, connection measurements, and traversal or path-finding. For more details, including some examples of what is included in each functionality category, please refer to the glossary (Annex I).

Of the 139 entries, 31 products listed some form of network analysis functionality. Table F-6 shows the portion of these products capable of doing each type of network analysis as input in the database.

Table F-6: The Network Analysis Functionalities of the 31 Products for which they were Specified in the Survey Database

Type of Analysis	Number of Products
Centrality measures	20
Cluster recognition	17
Connection measures	23
Traversal or path-finding	23

There were 10 products found that had all 4 of the network analysis functionality groups: [20], [23], [24], [30], [31], [32], [33], [34], [35], and [36].

F.1.3.2 Visual Abstraction

The additional visual presentation methods found in 12 products were: line charts; pie charts; area charts; bar charts; scatter plots; x-y plots.

F.1.3.3 General Analysis

Of 139 products, 19 (14%) included additional analysis methods.

F.1.4 Visual Enhancements

The "Visual Enhancements" main category lists various visualization methods that can enhance the user's understanding of the network. For example, animation can help to illustrate how the network changes over time. An immersive environment can add cues via audio and haptics. Information can be overlaid on a map.

Of 139 entries, 17 listed some form of enhancement. The majority (13) indicated animation as a feature: time-evolution of the network is featured in 6 products: [37], [29], [38], [39], [19] and [40]. In 1 of these products, the transition while zooming is animated; the remainder animate the transition between layouts. Distortion (e.g. fish-eye, 3-D hyperbolic projection) was listed for 4 products, including [5] and [28]. One product offered multiple related views of the networks [41].

F.1.5 User Interaction

The "User Interaction" main category lists the various ways that a user may interact with the software. Seventy-one products had entries for user interactions:

28	drag & drop	21	reposition	20	cut & paste	20	undo/redo
16	select	12	rotate	11	layers	10	drill down
9	resize	7	filter	7	scroll	3	search
3	sensory	3	spreadsheet	2	command line		

F.1.6 Deployment

The "Deployment" main category includes issues that the user may need to consider when deploying the software. This includes:

- **Type**: The software is stand-alone, Web-based, or components for building tools.
- **Platform**: The operating systems the software may run on.
- **Extensibility**: The languages used if the software is an API and/or can be otherwise extended or modified in some way.
- **Interoperability**: The methods by which the software may interact with other external software (import/export of industry standard file formats, remote procedure calls, etc.).
- **Scalability**: An indication of how well the software and its associated algorithms scale to large datasets.
- **Hardware**: Any required specialized hardware.
- **Users**: Does this software support just one user or can it be used in a multiuser environment? Can the software be networked?
- **Status**: The availability of the software, e.g. whether the software is commercially available, in development, no longer supported, or a research prototype without a public release.
- **Cost**: A rough indication of how expensive the software is.

F.1.6.1 Type

The Deployment sub-category "Type" includes a field to indicate whether a product was open source. This is a questionable placement of this property, and as such the open source component was analysed separately.

Of 128 products, 92 (73%) were stand-alone applications, 10 of which are also capable of being used as components to build other tools, and 14 of which were Web-based. Twenty-seven percent were exclusively intended to be used to build applications.

Assuming that products with a deployment type specified were closed-source if not explicitly marked open source, 48 of 132 products (36%) were open-source.

F.1.6.2 Platform

Of the 105 products that specified their supported platforms, the platform-independent Java language was the most prominent at 40% of the products using that language. 22% were developed solely for Windows, while only 3% were developed solely for Linux. Two use a proprietary hardware device, and 1 was developed for the PocketPC. The remaining products have been ported to various combinations of Windows, Linux, UNIX and MacOS platforms.

F.1.6.3 Extensibility

The languages identified for extending the functionality of a product were:

- .NET
- ActiveX
- BeanShell
- C
- C#
- C++
- CGI
- COM
- Java
- JavaScript
- JRuby
- JSP
- MFC
- Perl
- PHP
- Python
- Tcl/Tk
- VBS
- Visual Basic
- XML

Seventy-three of 139 products (52.5%) were identified as extensible; more then half of these used Java.

F.1.6.4 Interoperability

Based on the text field entries, the most commonly used formats for import and export are, with the number of products in parentheses: GraphML (12), GML (9), Pajek (7), and dot (3). Other data formats are listed, but are not standardized, such as CSV (comma-separated values), XML, and Oracle or MySQL database tables. The products listing these for interoperability are only interoperable to the extent that data can be manipulated before or after input or output.

F.1.6.5 Scalability

This field is intended to give an indication of how well the software and its associated algorithms scale to large datasets. If the link/node scalability is “Unlimited”, this is understood to mean that the size of the datasets that the software can handle is only limited by the CPU and memory of the computer on which the software is operating. Note that this field does not give an indication of how aesthetically pleasing a layout is for large networks, only whether the software is capable of processing large networks in reasonable time scales.

The survey indicates that the scalability of nodes and links are tightly coupled, i.e. the maximum number of links scales approximately with the maximum number of nodes. For 80% of the entries, however, the scalability is unknown. Twenty-two products claim unlimited scalability, and the remaining 7 products indicate being limited to less than 100000 nodes.

F.1.6.6 Hardware

Specialised hardware, i.e. hardware that may not be commonly found on a workstation, may include a 3-D graphics accelerator, a data glove (for haptic interaction), an electronic whiteboard, a graphics tablet, a joystick, a large screen display, or a virtual reality headset.

Of 139 products, only 5 required specialised hardware: 4 required a 3-D graphics accelerator and 1 required an electronic whiteboard with click and drag.

F.1.6.7 Users

Few products had the user deployment listed. Nineteen allowed both multiple users and networked users, 10 were single-user and 1 was multiple user, but not networked.

F.1.6.8 Availability

Of the 104 products for which the status was specified, 39% were commercially available. Some of these (3%) were also available in freeware or shareware versions. Overall, 43% of the products were freeware. Twenty-four percent were research prototypes, only 2% of which were listed as In-house Use (unshared).

F.1.6.9 Cost

For 53 of the 139 products, the cost is marked "unknown" in the survey. The cost of the remaining products is shown in Table F-7.

Table F-7: The Distribution of Purchase Price of Software Packages. There were 86 products with a specified value. All cost data is in U.S. dollars unless otherwise noted.

Number	% of Products	Cost Category
45	32.37	Free
12	8.63	$101 – $1000
10	7.19	$1001 – $5000
8	5.76	Free for non-commercial use
4	2.88	$5001+
3	2.16	Complicated
2	1.44	Free for academic use
2	1.44	$1 – $100

F.2 LITERATURE SURVEY STATISTICS

A survey of the current research in network visualisation was performed to assess the current state, trends, and future directions. The survey focussed on articles specifically related to network visualisation. Since we are interested in recent trends, only those articles published after 1999 were considered. Each article was examined and assigned keywords as descriptors of the focus of the work. The keywords and their meanings are shown in the table below.

Table F-8: The Categories, Keywords and Keyword Definitions used in Classifying the Network Visualisation Articles Collected for the Research Paper Survey

Category	*Keyword*	*Definition*
Theory	Framework	The article presents a visualisation framework or taxonomy.
	Evaluation	The article includes an evaluation of a system or methodology.
	Survey	The article presents a survey of previous work in visualisation.
Node Representation	New layout	A novel layout algorithm is presented.
	Modified layout	A layout which is a modification or improvement of an existing method is presented.
	Hybrid layout	A new layout is presented that combines two or more existing layout algorithms or methodologies.
	Simultaneous representation of data sets that share vertices	Layouts are investigated to simultaneously display two networks that share common vertices.
	Linked representations	Two or more views of the data are displayed simultaneously, where selection of an item on one view leads to selection of the same items on the other views.
	Icons	Icons are discussed in the article.
Dimensionality	3-D	Three-dimensional displays are used.
	Overlaying data	Information is overlaid on a network.
	Immersive environment	An immersive environment is discussed.
Scalability	Large	The large data set problem is addressed in some way.
	Reduction	A method of data reduction is presented.
	Clustering	A method of clustering data is presented.
	Space optimization	A method for optimizing the use of screen space is presented.
	Small screen	The special case of small screen displays is discussed (i.e. scaling *down*).
Human Aspects	Interactive	The method or system described incorporates user interaction.
	Mental map preservation	The article discusses the end-user's need to preserve the mental map when changes are made to the diagram.
	Animated change of focus	Animation is used in user interaction to preserve the mental map.
Specific Type of Network	Highly-connected	Most nodes are connected to most other nodes.
	Scale-free	Most nodes are connected to only one other node.
	Small-world	You can reach any node from any other node in less than 7 steps.
	Application-specific	The article presents a solution to a problem in a specific application area.
	System	The article presents a system or software package.
Dynamic	Dynamic	Attention is paid to the temporal evolution of the network.
	Animation	Layout animation is employed to show temporal evolution.

Research on network visualisation can be found in publications related to both information visualisation and graph drawing. Papers from the annual IEEE Symposium on Information Visualization (InfoVis) conference were studied in-depth, forming more than 50% of the entries. Other entries were found via Internet searches and IEEE/ACM searches of the years 2000 – 2007. Table F-9 lists the journal and conferences from which the literature was collected.

Table F-9: Journals and Conferences from which the Literature was Harvested

Visualisation Journals

- IEEE Transactions on Visualization and Computer Graphics
- Information Visualization

Visualisation and HCI Conferences

- IEEE Symposium on Information Visualization (InfoVis)
- Graph Drawing Symposium (GD)
- EUROGRAPHICS
- INFOVIS AUSTRALIA
- The SIGCHI Conference on Human Factors in Computing Systems
- Visualization for Computer Security (VisSEC)
- APVIS
- AVI
- BELIV AVI WORKSHOP
- Symposium on Visualisation
- IEEE Visualization
- ACM Virtual Reality Software and Technology
- ACM SIGGRAPH International Conference on Virtual Reality Continuum and its Applications
- Pan-Sydney area workshop on Visual information processing
- International conference on Human computer interaction with mobile devices and services
- International Conference on Intelligent User Interfaces
- Proceedings of the conference on Information Visualization

Application-Area Journals and Magazines

- Connections
- ACM Transactions on Information Systems
- American Journal of Sociology
- Journal of Computing Sciences in Colleges
- Journal of Social Structure
- Bioinformatics Journal
- IEEE Computer
- Discover Magazine

Application-Area Conferences

- SPIE Conference on Visualization and Data Analysis
- IIG Workshop on Networks, Management and New Governance
- American Statistical Association Section on Statistical Graphics
- West Point Information Assurance
- Knowledge and Data Discovery (KDD)
- Webometrics, Informetrics and Scientometrics
- ACM Symposium on Software Visualization
- Software Visualization
- Parallel and Distributed Computing, Applications and Technologies
- Australasian conference on Computer science
- IEEE/WIC/ACM International Conference on Web Intelligence
- ACM Symposium on Applied computing
- IEEE International Conference on Networks
- ACM/IEEE-CS joint conference on Digital libraries

F.2.1 Current State of the Research

According to Stephen Eick, who has been involved with the annual Information Visualization (InfoVis) conference since its inception, the state of information visualisation research as a whole has improved greatly since 1995 [42]. The number and quality of papers has increased, and the topics are stable. The network-specific papers collected in this survey also reflect this trend. The number of network-specific articles found in total increases steadily from 2000. Of the papers presented at the InfoVis conferences, the portion of articles focussing on networks has also increased since 2000. The data for this trend is shown in the table below.

Table F-10: The Number of Network-Related Papers Found, and the Portion of Network-Related Papers Presented at the InfoVis Conference

Year	*Total Network Papers*	*Network Papers at InfoVis*	*Total Papers at InfoVis*	*Portion of Network Papers at InfoVis*
2000	9	6	26	23%
2001	11	10	34	29%
2002	14	9	23	39%
2003	13	7	29	24%
2004	19	14	29	48%
2005	25	14	31	45%
2006	28	10	23	43%

Figure F-3 shows that the total number of articles in node placement, scalability and theory has increased since 2000. Figure F-4, however, shows the same data relative to the total number of papers collected. Relatively speaking, these three categories have actually remained rather constant. The growth seen in Figure F-3 is due to the overall growth of the information visualisation field. What does stand out, however, is the appearance of literature on dynamic networks in 2002. Not evident in these figures, due to the grouping of keywords into categories, is the recent appearance of literature on small screen representations, to be expected with the recent popularity of PDAs.

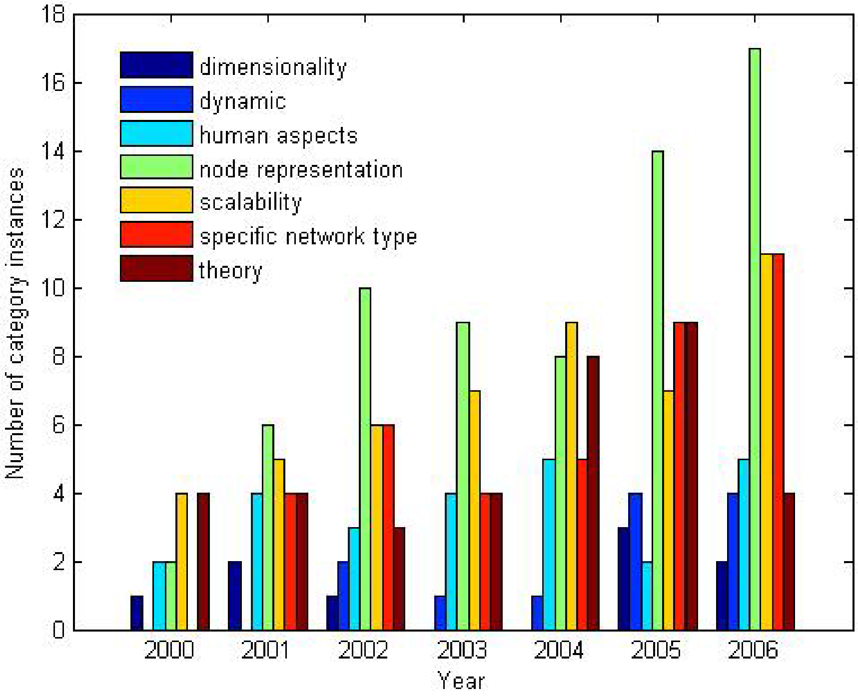


Figure F-3: The Number of Instances of a Category for Each Year from 2000 to 2006.

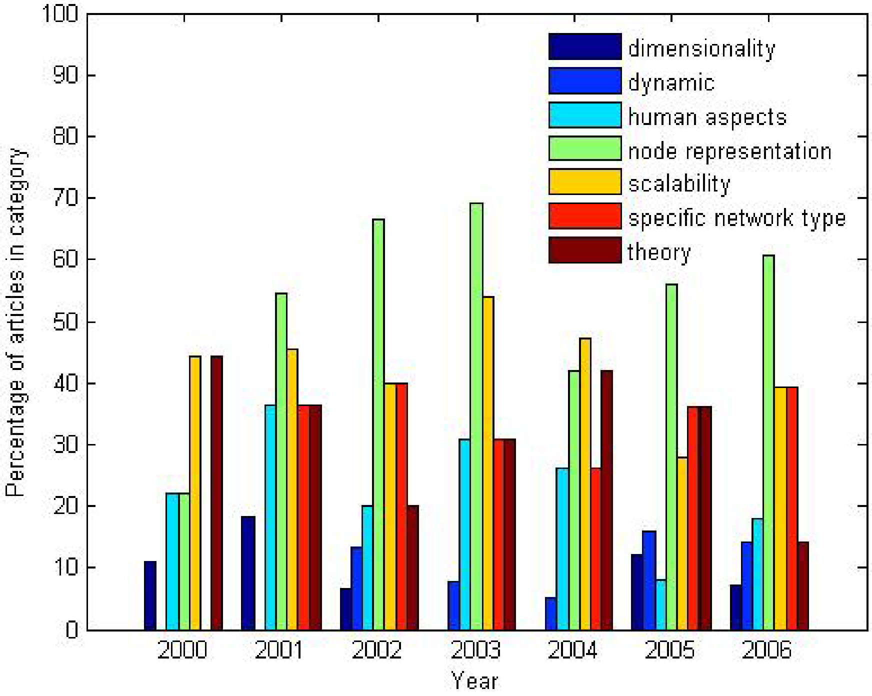


Figure F-4: The Percentage of Articles with Each Keyword from 2000 to 2006.

In Table F-11, it can be seen that the majority of the papers were generic in nature. Social networks are the next highest in total number collected, suddenly rising in 2005. The visualisation of online social networks behaves in the same manner, but with lesser total numbers. Biological sciences provide the most network visualisation articles from the sciences, with a fairly steady count over the period. Articles regarding visualisation of the Internet and computer networks began appearing in 2003 and have remained at about the same levels since.

Table F-11: The Trends Seen in the Field of Study to which the Content of the Paper was Applied. Note that very few papers were collected in 2007, since the data was collected early in the year.

Field of Study	*2000*	*2001*	*2002*	*2003*	*2004*	*2005*	*2006*	*2007*	*Total*
Generic	9	9	10	7	13	15	17	1	*81*
Social Networks	0	0	2	0	0	4	5	2	*13*
Biology	0	2	2	1	0	1	2	1	*9*
Computer Networks and the Internet	0	0	0	3	1	3	2	0	*9*
Online Social Networks	0	0	0	0	0	1	2	2	*5*
Library Science	0	0	1	0	2	0	0	0	*3*
Software Development	0	0	1	0	1	0	1	0	*3*
State Transition Diagrams	0	1	0	0	0	0	1	0	*2*
Genealogy	0	0	0	0	0	1	0	0	*1*
Engineering and Physical Sciences	0	0	0	0	0	1	0	0	*1*

F.3 REFERENCES

[1] Gort, A. and Gort, J., (2007), "Survey of Network Visualization Tools", *DRDC Ottawa CR 2007-280*, Defence R&D Canada – Ottawa.

[2] Auber, D., (2008), "Tulip Software" (online), http://tulip-software.org (Access Date: 14 February 2008).

[3] Oreas Optimization Research and Software, (2008), "GoVisual Automatic Diagram Layout Software Libraries" (online), http://www.oreas.com (Access Date: 12 June 2008).

[4] Robertson, G.G., Mackinlay, J.D. and Card, S.K., (1991), “Cone Trees: Animated 3D Visualizations of Hierarchical Information”, In *CHI ‘91: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 189-194, New York, NY, USA: ACM.

[5] AbsInt Angewandte Informatik GmbH, (2007), “aiSee Graph Layout Software” (online), http://www.aisee.com (Access Date: 12 June 2008).

[6] Biuk-Aghai, R.P., (2006), “Visualizing Co-Authorship Networks in Online Wikipedia”, In *Communications and Information Technologies (ISCIT)*, pp. 737-742.

[7] Herman, I., Melançon, G. and Marshall, M.S., (2000), “Graph Visualization and Navigation in Information Visualization: A Survey”, *IEEE Transactions on Visualization and Computer Graphics*, 6(1), 24-43.

[8] Visual Analytics, Inc., (2008), “VisuaLinks Description at a Glance” (online), http://www.visualanalytics.com/products/visualinks/ (Access Date: 12 June 2008).

[9] Di Battista, G., Eades, P., Tamassia, R. and Tollis, I.G., (1994), “Algorithms for Drawing Graphs: An Annotated Bibliography”, *Comput. Geom. Theory Appl.*, 4(5), 235-282.

[10] Fruchterman, T.M.J. and Reingold, E.M., (1991), “Graph Drawing by Force-Directed Placement”, *Softw. Pract. Exper.*, 21(11), 1129-1164.

[11] Noack, A., (2003), “An Energy Model for Visual Graph Clustering”, In *Proceedings of the 11th International Symposium on Graph Drawing (GD 2003)*, pp. 425-436, Berlin: Springer-Verlag.

[12] Kamada, T. and Kawai, S., (1989), “An Algorithm for Drawing General Undirected Graphs”, *Information Processing Letters*, 31(1), 7-15.

[13] Frick, A., Ludwig, A. and Mehldau, H., (1995), “A Fast Adaptive Layout Algorithm for Undirected Graphs”, In *GD ‘94: Proceedings of the DIMACS International Workshop on Graph Drawing*, pp. 388-403, London, UK, Springer-Verlag.

[14] Koren, Y., Carmel, L. and Harel, D., (2002), “ACE: A Fast Multiscale Eigenvectors Computation for Drawing Huge Graphs”, In *INFOVIS ‘02: Proceedings of the IEEE Symposium on Information Visualization (InfoVis’02)*, p. 137, Washington, DC, USA, IEEE Computer Society.

[15] Harel, D. and Koren, Y., (2000), “A Fast Multi-Scale Method for Drawing Large Graphs”, In Marks, J., (Ed.), *GD 2000: 8th International Symposium on Graph Drawing*, Lecture Notes in Computer Science, pp. 183-196, Springer.

[16] University of Lethbridge, (2004), “GLuskap” (online), http://www.cs.uleth.ca/vpak/gluskap/ (Access Date: 12 June 2008).

[17] Daisy Analysis Ltd., (2008), “Daisy Analysis” (online), http://www.daisy.co.uk/ (Access Date: 12 June 2008).

[18] Meyer, B., (1998), “Self-Organizing Graphs – A Neural Network Perspective of Graph Layout”, In *GD ‘98: Proceedings of the 6th International Symposium on Graph Drawing*, pp. 246-262, London, UK: Springer-Verlag.

[19] Wolfram Research Inc., (2008), "Introduction to Graph Drawing – Wolfram Mathematica" (online), http://reference.wolfram.com/mathematica/tutorial/GraphDrawingIntroduction (Access Date: 12 June 2008).

[20] Hagberg, A., Schult, D. and Swart, P., (2008), "NetworkX" (online), https://networkx.lanl.gov/wiki (Access Date: 12 June 2008).

[21] ILOG Inc., (2008), "ILOG JViews Diagrammer: Java and Ajax Diagrams and Dashboards" (online), http://www.ilog.com/products/jviews/diagrammer (Access Date: 19 February 2008).

[22] Dudek, M., (2008), "Zone Manager – Prefuse Zoned" (online), http://goosebumps4all.net/ZoneManager/ (Access Date: 12 June 2008).

[23] yWorks, (2008), "yFiles – Java Graph Layout and Visualization Library" (online), http://www.yworks.com/en/products_yfiles_practicalinfo_gallery.html (Access Date: 3 July 2008).

[24] Cyram Co., Ltd., (2008), "NetMiner Social Network Analysis Software" (online), http://www.netminer.com (Access Date: 3 July 2008).

[25] Tom Sawyer Software, (2008), "Tom Sawyer Software" (online), http://www.tomsawyer.com (Access Date: 12 June 2008).

[26] Universitt Bremen, (2005), "uDraw(Graph)" (online), http://www.informatik.uni-bremen.de/uDrawGraph/en/uDrawGraph/uDrawGraph.html (Access Date: 12 June 2008).

[27] Gansner, E., Koren, Y. and North, S., (2004), "Topological Fisheye Views for Visualizing Large Graphs", In *INFOVIS '04: Proceedings of the IEEE Symposium on Information Visualization*, pp. 175-182, Washington, DC, USA: IEEE Computer Society.

[28] Munzner, T., (2001), "H3Viewer" (online), http://graphics.stanford.edu/~munzner/h3/ (Access Date: 12 June 2008).

[29] Forrester, D., Kobourov, S.G., Navabi, A., Wampler, K. and Yee, G. (2005), "Graphael" (online), University of Arizona, http://graphael.cs.arizona.edu (Access Date: 13 June 2008).

[30] O'Madadhain, J., Fisher, D., Nelson, T., White, S. and Boey, Y.-B., (2008), "JUNG – Java Universal Network/Graph Framework" (online), http://jung.sourceforge.net (Access Date: 19 February 2008).

[31] orgnet.com, (2008), "InFlow Social Network Analysis Software & Services for Business, Communities, and Government" (online), http://www.orgnet.com/inflow3.html (Access Date: 13 June 2008).

[32] Batagelj, V., Mrvar, A. and Zavernik, M., (2008), "Pajek Wiki" (online), http://pajek.imfm.si/ (Access Date: 13 June 2008).

[33] Analytic Technologies, (2008), "Social Network Analysis & Cultural Domain Analysis" (online), http://www.analytictech.com/ (Access Date: 13 June 2008).

[34] Dwyer, T., Eckersley, P. and Cowling, J., (2004), "Wilmascope 3D Graph Visualisation System" (online), http://wilma.sourceforge.net/ (Access Date: 13 June 2008).

[35] Boost.org, (2008), "Boost C++ Libraries" (online), http://www.boost.org/ (Access Date: 13 June 2008).

[36] Csrdi, G. and Nepusz, T., (2008), "The igraph Library for Complex Network Research" (online), http://cneurocvs.rmki.kfki.hu/igraph/ (Access Date: 13 June 2008).

[37] CASOS, (2008), "DyNet Software" (online), Carnegie Mellon University, http://www.casos.cs.cmu.edu/projects/DyNet/ (Access Date: 13 June 2008).

[38] Stanford University, (2008), "Social Network Image Animator" (online), Stanford University, http://www.stanford.edu/group/sonia (Access Date: 13 June 2008).

[39] MIT Center for Collective Intelligence, (2008), "iCKN – Innovative Collaborative Knowledge Networks" (online), MIT, http://www.ickn.org/ (Access Date: 13 June 2008).

[40] Kobourov, S., (2006), "TGRIP: Temporal Graph dRawing with Intelligent Placement" (online), http://tgrip.cs.arizona.edu/ (Access Date: Not available as of 13 June 2008).

[41] Le, V. and Navabi, A., (2003), "Simultaneous Graph Drawing: Layout Algorithms and Visualization Schemes" (online), http://simg.cs.arizona.edu/ (Access Date: 13 June 2008).

[42] Eick, S.G., (2005), "Information Visualization at 10", *IEEE Computer Graphics and Applications*, 25(1), 12-14.

Annex G – CHARACTERISATION OF NETWORK VISUALISATION APPROACHES WITH RM-VIS

A. Bouchard and R. Vernik

G.1 INTRODUCTION

In recent years, surveys of Information and Network Visualisation technologies have been performed (e.g. Bouchard [2], Herman et al [6]), identifying over 70 different products to facilitate the discovery of Network Visualisation technologies. However most of these surveys use an *ad hoc* reference model to characterise the surveyed products and suggest potential technologies for the particular application that required a survey (e.g. intelligence, counterterrorism, banking). There is clearly a need to characterise network visualisation technologies with an appropriate visualisation reference model.

Several taxonomies have been developed to help characterise and describe Information Visualisation technologies such as provided by Card et al [3]. However, these typically focus on the visualisation approach without reference to the context within which it is used. The Information Visualisation Action Group (AG3) of The Technical Cooperation Program (TTCP) developed a reference model (cf. [4][5][6]) for visualisation (RM-Vis) to characterise and showcase visualisation approaches within their usage contexts. The use of a reference model such as RM-Vis allows the rapid identification of relevant approaches for particular activities, as well as providing support for evaluation and transitioning of the approaches in operational environments.

This Annex first introduces the TTCP RM-Vis reference model and describes how it has been used for characterising network visualisations technologies in terms of domains of use, the descriptive aspects (i.e. what they describe) and the approaches that they use for presenting the information. We discuss our experiences in using these approaches for the development and use of a system called C2NetVis which characterises and showcases network visualisation technologies. The contexts of use and descriptive aspects of network visualisation approaches used in Command and Control environments are provided.

Finally this Annex discusses how we are extending on the work done by TTCP as part of Project Imago. This project is developing a Web-based distributed environment that can be used to collaboratively define, prototype, evaluate, and transition visualisation approaches for C2.

G.2 REFERENCE MODEL FOR VISUALISATION

The TTCP Action Group on Information Visualisation developed a Reference Model framework (cf. [4]) for the application of Visualisation approaches (RM-Vis), which has been used to support the characterisation, identification and showcasing of visualisation approaches in several domains including C3I. This framework allows visualisation solutions to be defined in terms of their context of use, the representation and presentation techniques used, and key features of tool support provided such as types of user interactions and deployment support.

As shown in Figure G-1, RM-Vis has three key dimensions:

- The **Domain Context** is a model that defines the focus for the application of visualisation approaches i.e. ***where*** visualisation approaches will be applied, ***who*** will be supported, and ***why*** the approaches are needed.

- **Descriptive Aspects** (DA) define what needs to be described for particular domain contexts. For example, DAs could be defined in terms of the various elements (or things) that are of importance, the relationships between those elements and particular attributes that describe the elements and relationships.
- The **Visualisation Approach** dimension defines how the required information can be provided through computer-based visualisation. Approaches are characterised in terms of the visual representations used (e.g. graphs, charts, maps), visual enhancements (e.g. use of overlays, distortion, animation), interaction (direct manipulation, drag and drop, haptic techniques, etc.), and deployment which includes the computing environment (display devices, COTS software) and advanced deployment techniques such as intelligent user support and enterprise integration.

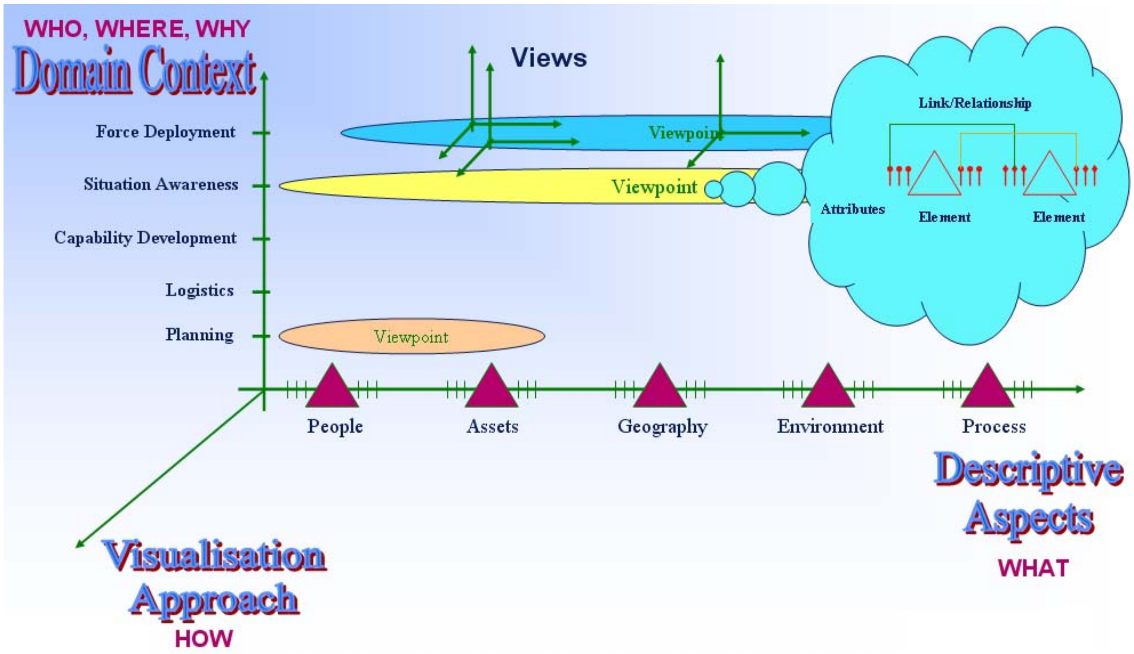


Figure G-1: RM-Vis Framework.

In parallel to the development of the reference model, the members of AG-3 created three instantiations of a database containing views referencing the model. C3I-Vis, MIL-Vis, and G-Vis were created to characterise and showcase visualisation approaches in the C3I, Military, and general domains. This Annex discusses how these approaches have been used in the development and use of a system called C2NetVis which characterises and showcases network visualisation technologies.

G.3 CHARACTERISATION OF NETWORK VISUALISATION APPROACHES

A fourth instantiation of the reference model database has been developed, called C2NetVis, to contain Network Visualisation approaches for the Command and Control domain context. The characterisation of the visualisation approaches required the instantiation of the three main dimensions of the reference model: domain context, descriptive aspect, and visualisation approach. From the two first dimensions have been derived a list of viewpoints, which corresponds to the main tasks requiring network visualisations in C2.

Finally 55 network visualisations have been characterised using the instantiation of the model. The following sections present these results.

G.3.1 Domain Context

The Domain Context (DC) is a model, which defines the focus for the application of visualisation approaches i.e. ***where*** visualisation approaches will be applied, ***who*** will be supported, and ***why*** it is needed. A domain context can be generated from existing enterprise models and tailored for the particular application of the reference model (cf. [4]). For example, various tasks require visualisation approaches in support to air operations, which tasks might be defined in term of roles and activities in the model.

This document focuses on the use of network visualisations in the Command and Control context. The same visualisations approaches could potentially be used in many other fields such as Health and Finance, but the pertinence and efficiency of these visualisations need to be assessed in these particular contexts. There are many reasons why a visualisation might be appropriate to achieve a task in one context but inappropriate for a similar tasks in a different context. Historical, technical, and even social aspects may influence the adoption or rejection of visualisations.

The Command and Control domain inherits its specificities from the more general military domain, which has a standardised way of categorising tasks based on pre-defined doctrines and procedures. The instantiation of the model follows this categorisation. Giving a particular task in the military domain, the where, who, and why aspects of the task are generally definite and have been used to define the model. The C2NetVis reference model adopted the domain context model from a previous instance of the RM-Vis database, C3I-Vis, with only minor changes. Table G-1 outlines the main aspects of the domain context model used to characterise the network visualisations.

Table G-1: Domain Context Model

	Category	Attribute
Where	Level of Command	• Operational, Strategic, Tactical
	Environment	• Air, Land, Maritime, Joint, Littoral, Space, Urban
	Area	• Acquisition, Communications, Development, Engineering, Intelligence, Operations, Personnel, Plans, Requirements, Research, Training
	Scenario	• Humanitarian Assistance, Low/Medium/High Intensity Conflict, Peace Support, Special Ops
Who	Role	• COS, Commander, J2, J6, J7, J8, Intel Analyst, Logistics Officer, Ops Officer, Support Engineer
Why	Activity	• Analysis, Assess, Assign, Execute, Monitor, Plan, Report, Schedule, Track

In order to keep the model simple, a hierarchical definition has been adopted. However a more comprehensive model would have to be defined as a domain ontology, including relationships and constraints between categories. For example, certain roles and areas seem to have a close relationship (ex: intel analyst in intelligence area) while some others roles and areas are complete dichotomies (ex: Chief of Defence Staff in the Health or Finance domain).

G.3.2 Descriptive Aspects

Descriptive Aspects (DA) define ***what*** needs to be described for particular domain contexts. For example, DAs could be defined in terms of the various elements (or things) that are of importance, the relationships between those elements and particular attributes, which describe the elements, and relationships (cf. [4]). The DA dimension is a model by itself that is derived in the context of the DC model, identifying the entities that need to be represented in that context. The development of a DA model consists generally of using a more general model and adding domain specific aspects.

The DAs that are relevant to characterise network visualisation approaches in a C2 context is quite extensive mainly because, as mentioned earlier, the potential of use of network visualisations is extensive. As it was the case with the DC model, the C2NetVis reference model adopted the descriptive aspect model from the C3I-Vis reference model, with added network specific aspects. In fact, the core DAs in the C2 context is similar to the ones in many other contexts, many elements to be visualised are recurrent in many domains. For example the time, location, events, resource, and people aspects may be member of any DA model as these are common elements in visualisations. In the case of communication or computer networks, some more specific aspects such as computer, hardware devices, protocols, usage, and capacity need to be added to the model. In other physical networks, a road network for example, aspects as structure, speed, and weather need to be considered. Concerning social networks, other specific aspects such as identity, skills, influence, relationships, health, travel, and opinion are central elements displayed.

Again for the DA model a hierarchical approach has been selected to keep it as simple as possible and no particular relationships among the aspects has been established. In a more comprehensive model some relationships and constraints would be present. Some aspects may be meta-data of other aspects for example; a Currency descriptive aspect should always be attached to a Money aspect, a Time Zone to a Time aspect, and so on. Constraints should also be modelled. For example, two different aspects may represent opposite or sequential information and should be defined accordingly in the model. As an example a date of death is always greater or equal to a date of birth and reflects a change of state from alive to dead.

G.3.3 Viewpoints

A Viewpoint is a model of what needs to be described for particular domain contexts (cf. [4]). Specifically a viewpoint represents a task in a particular context, which requires the visualisation of different elements regardless of how this information is displayed. In other words the viewpoint lives in the two dimensional world of domain contexts and descriptive aspects. For example, an air traffic controller has the task, read viewpoint, to monitor the distribution of aircraft in space. The way by which the air controller achieves his task corresponds to the different approaches to visualise his viewpoint on screen.

As mentioned earlier, the number of tasks involving network analysis visualisations in C2 is extensive, as important is the number of viewpoints. The C2NetVis reference model characterises three viewpoints representing typical use of network analysis techniques. Table G-2 presents the viewpoints and their definition.

Table G-2: Example of Viewpoints

Viewpoint	Domain Context	Descriptive Aspect
Monitor Belligerent Activities	**Activity**: Analyse, Assess, Monitor, Track **Area**: Communications, Intelligence **Environment**: Joint, Land, Urban **Level of Command**: Strategic, Tactical **Role**: HQ J2, Intel Analyst **Scenario**: Special Ops, Low Intensity Conflict	**Communications**: Email, Phone **Events**: Sequence **Finance**: Currency, Money **Geography**: Area, City, Country, Origin, Destination, Location, Maps **Identity**: Name, Sex **Information**: Document, File, Opinion **Movement**: Flight, Travel **Occupation**: Activity, Engagement **Organisation**: Unit **People**: Belligerent, Group, Organisation, Warlord **Relationships**: Degree, Enemy, Friend, Non-friend **State**: Alive, Dead **Time**: Age, Critical, Current, Date, Duration, Interval **Transportation**: Vehicle, Car
Assess Robustness of Communication Network	**Activity**: Analyse, Assess **Area**: Communications **Environment**: ALL **Level of Command**: Tactical **Role**: Support Engineer **Scenario**: ALL	**Computer**: Hardware, Network **Geography**: Location, Maps, Latitude, Longitude **Telecommunication**: IP, Network, Parabolic-dish, Satellite **Usage**: Frequency
Team Building	**Activity**: Assign, Plan **Area**: Personnel, Training **Environment**: ALL **Level of Command**: ALL **Role**: Ops Officer **Scenario**: ALL	**Ability**: Skill **Assignment**: Mission, Order **Capacity**: Force **Events**: Scenario, Sequence **Identity**: Name, Sex **Occupation**: Activity, Engagement, Function, Jobs, Responsibility, Task, Work **Organisation**: Unit **People**: Group, Organisation, Person, Player, Soldier **Relationships**: Friend **State**: Ready, Standby, Not-Ready, Morale **Time**: Age, Deadline, Duration, Priority

G.3.4 Visualisation Approaches

The **Visualisation Approach** dimension defines ***how*** the required information can be provided through computer-based visualisation (cf. [4]). Approaches are characterised in terms of four independent sub-dimensions; the visual representation, visual enhancement, interaction, and deployment forming a visual abstraction in support to a viewpoint.

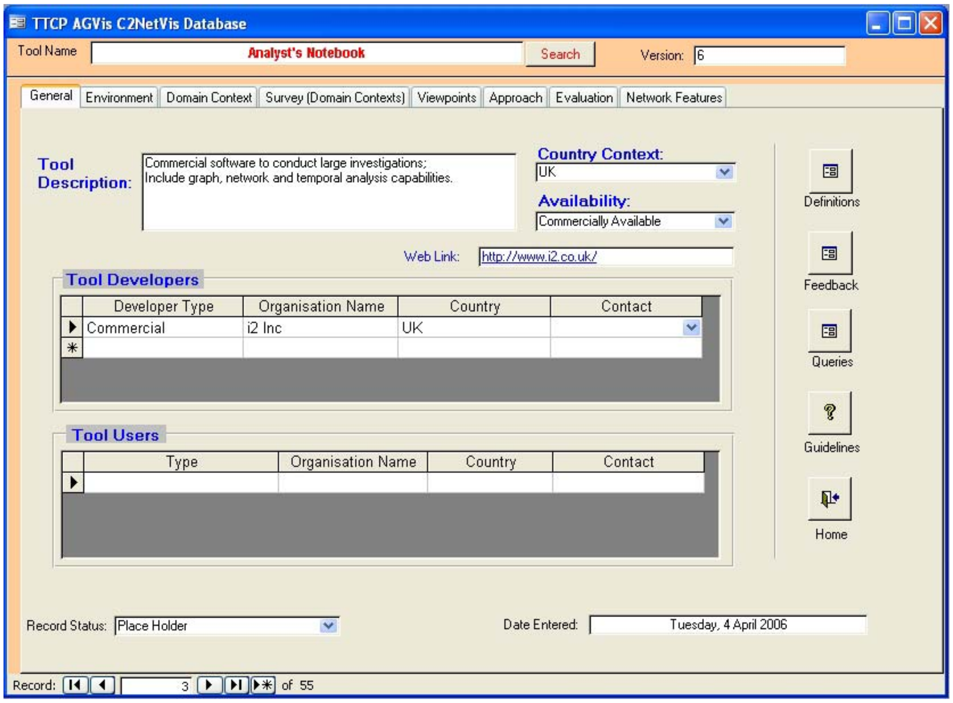
The C2NetVis database includes about two hundred different representations, enhancement, interaction, and deployment attributes. The representation sub-dimension refers to the techniques used in transforming data elements into visual forms. Card *et al* [2] taxonomy has been used to populate the representation dimension, which includes visual abstractions such as chart, colour, glyph, graph, icon, map, table, tree, etc. The enhancement dimension contains items that allow an improved presentation of the visual elements on screen using groupings, overlays, stereoscopy, distortion, animation, and others. As for the interaction dimension, it includes the techniques which allow a user to tailor visual information to specific needs, including various ways that the user can interact with the visual elements such as drag & drop, cut & paste, pan, resize, undo and redo, zoom. Finally the deployment dimension, although not an intrinsic part of a visual representation, it refers to those features which allow for the provision of cost effective visualisation solutions including the computing environment (display devices, COTS software, operating system, hardware) and advanced deployment techniques such as intelligent user support and enterprise integration.

The visualisation approach dimension is agnostic of the viewpoints it represents, therefore independent of the domain context and descriptive aspects, as it is only a way of rendering elements on screen and providing interaction capabilities. In other words, a set of descriptive aspects in a particular domain context can be represented using any combination of representation, enhancement, interaction, and deployment characteristics. Therefore the definition of the visualisation approach dimension is fairly stable and can be reused in various contexts.

G.3.5 Characterising Views

Characterising a view consists of positioning it into the three dimensional world of domain context, descriptive aspect, and visualisation approach as well as attaching a set of meta-data to the view, which might contain attributes such as a name, description, producer, and showcase examples. The first step in characterising a view consists of defining a viewpoint particular to that view, which contains the domain context and descriptive aspects information. Then the visualisation approach for that view may be characterised by selecting values from the four representation, enhancement, interaction, and deployment sub-dimensions. Finally, one can attach meta-data relevant to the view; attaching one or more showcase examples is particularly important so that the user can appreciate the effectiveness of the view in support to his viewpoints.

The C2NetVis database characterises 55 network analysis products. The characterisation of Analyst's Notebook geography view, as shown in Figure G-3, is presented below as a thorough example. Figure G-2 presents the meta-data for the view, as displayed in the general information tab of the database, and Figure G-3 presents the view itself.



Tool Name: Analyst's Notebook

Version: 6

Tool Description: Commercial software to conduct large investigations; Include graph, network and temporal analysis capabilities.

Country Context: UK

Availability: Commercially Available

Web Link: http://www.i2.co.uk/

Tool Developers

Developer Type	Organisation Name	Country	Contact
Commercial	i2 Inc	UK	

Tool Users

Type	Organisation Name	Country	Contact

Record Status: Place Holder

Date Entered: Tuesday, 4 April 2006

Record: 3 of 55

Figure G-2: C2NetVis Database Showing Analyst's Notebook Meta Data.

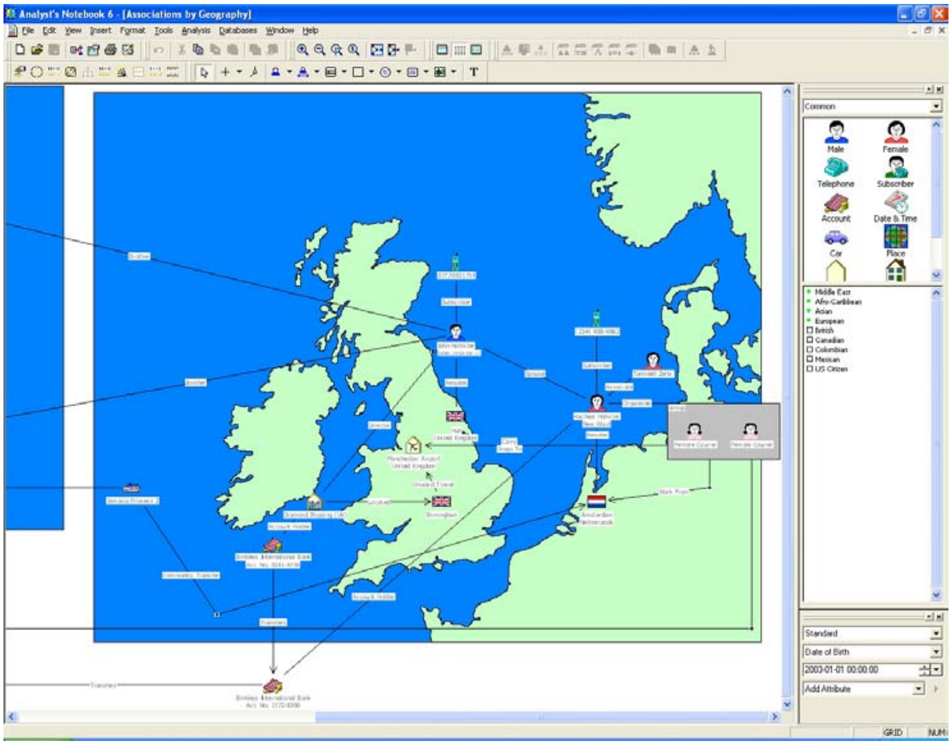


Figure G-3: Analyst's Notebook Geography View (Dataset Provided by i2 Inc.)

The definition of the viewpoint for that particular view is presented in Table G-3, referencing the view in the domain context and descriptive aspect dimensions.

Table G-3: Viewpoint for Analyst's Notebook Geography View

Domain Context	**Descriptive Aspect**
Activity: Analyse, Assess, Report, Schedule **Area**: Intelligence **Environment**: Joint, Land **Level of Command**: Strategic, Tactical **Role**: HQ J2, Intel Analyst **Scenario**: Special Ops, Low Intensity Conflict	**Communications**: Email, Phone **Family**: Brother, Sister **Finance**: Account, Money, Transfer **Geography**: Area, City, Country, Location, Maps **Identity**: Name, Sex, Flag, Nationality **Movement**: Flight, Travel **People**: Actor, Group **Possession**: Holder **Relationships**: Family **Time**: Age, Critical, Current, Date, Duration, Interval **Transportation**: Vehicle, Car, Ship

Looking at the Analyst's Notebook viewpoint, one can notice that many of its elements are also presents in the *Monitor belligerent activities* and *Team building* viewpoints, but fewer elements are shared with the *Assess robustness of communication network*. Therefore it can be assumed that this view might be appropriate to support these former tasks but might be less appropriate to support the latter. However it's not possible to assume the effectiveness of a view for a task unless a user actually evaluates the view in the context of the particular viewpoint. The subject of view evaluation is another complex subject by itself, although the closeness of a viewpoint to another one is a good indicator of effectiveness.

The viewpoint of the view being defined, the visualisation approaches supported by this view can be defined. Table G-4 presents the characterisation of the visualisation approaches for the Analyst's Notebook geography view.

Table G-4: Visualisation Approaches of Analyst's Notebook Geography View

Representation	Interaction
Figure Graph.directed Graph.layout.symmetric Graph.link Graph.link.text Graph.node Graph.node.nested Graph.node.icon Graph.node.text Map.2D Table List	Cut & Paste Drag & Drop GUI.Point and Click Highlight Pan Resize Scroll Select Undo & Redo Zoom
Enhancement	**Deployment**
Grouping Layering Overlay	Availability.Commercially available Extensibility.C# Extensibility.C++ Extensibility.COM Extensibility.COM+ Extensibility.NET OS.Windows.* Platform.PC.* Users.Multiple

G.3.6 Other Network Analysis Features

RM-Vis being a generic visualisation reference model, it fails to include some domain specific features of tools that might be important to record and search against. Some non-visual features are sometimes essential in the production of a visualisation although they can't be referenced in the model, an algorithm for example. In the particular example of network visualisations features such as analysis functions (shortest path, pattern analysis), type of networks, transformational and mathematical properties, and constraints would be of interest to model.

For sake of comprehensiveness, the C2NetVis database includes a list of non-visual features that visualisation approaches can reference. These extra features are attached as meta-data to the views and can then be search as other meta-data entities in the database.

G.3.7 Using the Database

The database has been populated with products surveyed from different studies and showcase examples of views from these products have been characterised using the framework. There are many ways of interrogating the database, listing all the views being the most straightforward. Predefined queries are also available for most common requests, such as listing views by domain context, and more complex interrogations, such as views that support intelligence in a peace keeping operation. Other queries can be defined by the user.

As an example of query, the database has been searched for views supporting the viewpoints defined in Table G-5. The following table shows the result of the query.

Table G-5: Query Result of Views by Viewpoints

Viewpoint	View	
Monitor Belligerent Activities	Analyst's Notebook Daisy InFlow NetMap	NetMiner Starlight VisuaLinks
Assess Robustness of Communication Network	Analyst's Notebook Daisy FATCAT NetMap	NetMiner Starlight VisuaLinks
Team Building	Agna Analyst's Notebook InFlow KrackPlot MultiNet Negopy Netdraw NetMap NetMiner	Netvis Pajek SocioMetrica Starlight StOCNET UCINET Visone VisuaLinks

These results clearly indicate that some more generic views, such as Analyst's Notebook, Starlight, VisuaLinks, might be applied for various purposes while others are more specialised and then less applicable for certain tasks. Also it can be noticed that all the views used in the context of monitoring belligerent activities can also be used in a team building process. This might be explained by the fact that, although the task of monitoring belligerent activities does not involve building a team as a sub-task, the viewpoint associated to the task of building a team is probably partially or totally included within the viewpoint of monitoring belligerent activities. In that case the inner viewpoint shares the same views of the including other.

Also, it has to be noted that software libraries might potentially be used in the context of all viewpoints, depending on the capabilities of the library but also on the host product using the library. Therefore no libraries have been included in the result of the query for clarity purposes.

Finally, even though scientists and analysts may populate and navigate the C2NetVis database itself, and similarly the other three database instantiations, the main interest in the database is the collaboration among many individuals to share approaches, evaluations, and for showcasing purposes. As the number of users and operations on the database increases it becomes difficult to synchronise remotely located instances. An ideal configuration would consist of a unique database instance shared by many users, distributed over an enterprise bus. The Imago project, discussed in Section G.4, is addressing this issue.

G.4 FUTURE WORK

Imago is a project being lead by the Defence Science and Technology Organisation (DSTO) to support the development, evaluation and transitioning of Information Visualisation approaches for Command and Control (C2). Imago is a distributed environment that can be used to rapidly prototype, evaluate, and transition visualisation approaches for C2. The platform will provide a means of integrating and sharing the output of visualisation tools, storing, accessing and managing showcase examples of visual representations via an underlying reference model, and providing access to underlying data sources provided through simulation, representative data, and/or operational data.

The work done in characterising network visualisations in terms of their usage contexts will be uploaded into Imago as one of its initial knowledge bases. Imago uses a semantic network modelling approach to support various queries and reasoning about the use of visualisation approaches within target domain contexts. The inference engine running in the back-end will allow querying and discovering complex relationships amongst data in the various model dimensions. The distributed aspect of Imago will allow multiple users to collaborate and share their knowledge of the effectiveness of views for their tasks. An evaluation framework is a core element of the Imago system. This facility will support the evaluation of visualisation approaches within actual usage contexts to provide information which will aid more effective transitioning into practice and for the development of enhanced visualisation techniques.

G.5 CONCLUSION

This Annex introduced C2NetVis, the instantiation of a visualisation reference model and a database that characterises network visualisation approaches for the Command and Control domain. C2NetVis falls within a larger research program aiming to develop, evaluate, and transition visualisation approaches based on the RM-Vis reference model. Previous work involved the development of G-Vis, C3I-Vis, and MIL-Vis, three other instantiations of the RM-Vis model. C2NetVis serves as a knowledge base for the discovery, evaluation, and transition of network visualisation approaches. The database characterises these approaches using the three main dimensions of the reference model; domain context, descriptive aspect, and visualisation approach. C2NetVis also defines viewpoints, which are models of what needs to be described for particular domain contexts and usually correspond to user’s tasks.

A typical use of the database consists of firstly creating viewpoints corresponding to the network visualisation tasks to be achieved in a C2 context. The relevant views can be discovered by searching the database against the predefined viewpoints. Finally the effectiveness of views can be evaluated in their context of use.

The main interest in C2NetVis is the collaboration between various users and scientists to share their experience on using network visualisations in Command and Control. The resulting database will serve as an initial knowledge base of the Imago system, which will extend the current system by providing a Web-based distributed environment that can be used to collaboratively define, prototype, evaluate, and transition visualisation approaches for C2.

G.6 REFERENCES

[1] Bouchard, A., (2004), “Link Analysis & Visualization Product Review”, *DRDC Valcartier TM-2004-175*, Defence R&D Canada – Valcartier.

[2] Card, S.K., Mackinlay, J.D. and Shneiderman, B., (1999), "Readings in Information Visualization: Using Vision to Think", Morgan Kaufmann Publishers.

[3] Gouin, D., Evdokiou, P. and Vernik, R., (2002), "A Showcase of Visualization Approaches for Military Decision Makers", In *NATO's Research & Technology Organisation (RTO) Information Systems Technology Panel Workshop: Massive Military Data Fusion and Visualisation: Users Talk with Developers*, Halden, Norway.

[4] Gouin, D., Vernik, R.J., Evdokiou, P., Houghton, P., Barnes, M., Monk, D. and Rosenblum, L., (2002), "Information Visualization in C3I: Final Report of TTCP C3I AG-3", *DOC-C3I-AG3-1-2002*, TTCP C3I AG-3.

[5] Herman, I., Melançon, G. and Marshall, M.S., (2000), "Graph Visualization and Navigation in Information Visualization: A Survey", *IEEE Transactions on Visualization and Computer Graphics*, 6(1), 24-43.

[6] Vernik, R., (2000), "A Proposed Reference Model Framework for the Application of Computer-Based Visualisation Approaches", Presentation at the *NATO Workshop on Visualisation of Massive Military Datasets*, Québec, Canada.

Annex H – THE VISTG REFERENCE MODEL

M.M. Taylor

H.1 INTRODUCTION

The VisTG Reference Model has a long history, having been initially proposed by DRG Panel 3 RSG-30, and described more completely in several chapters of the Final Report of IST-013/RTG-002 [1]. This Annex extends the IST-013 description and relates the VisTG Reference Model to other proposed frameworks for visualisation and for the human interface to visualisation systems. As such, it does not refer specifically to the visualisation of networks. The specifics of network representation fit within the framework of the VisTG Reference Model, but are not intrinsic to it.

The VisTG Reference Model was initially inspired by W.T. Powers' "Perceptual Control Theory" of psychology (PCT [2]). The fundamental idea behind PCT has been known at least since the time of Aristotle, and probably much longer: people act so as to get what they want, in the face of unpredictable events in the world in which they live. PCT refines this notion, using technical ideas developed in the 20th century.

According to PCT, all deliberate actions are performed in order to affect something perceptible about the world, and in particular, to bring one's perception of some aspect of the world closer to the way one would wish it to be, perhaps in the face of external forces that might alter it in other ways. This very simple, and on reflection necessarily true [3], statement has profound consequences if it is taken literally. We will not follow those consequences here; Powers suggested many of them in his 1973 book. Here, all we need to note is that the basic statement implies the existence of control systems in the engineering sense. Figure H-1 shows one such elementary control unit (ECU).

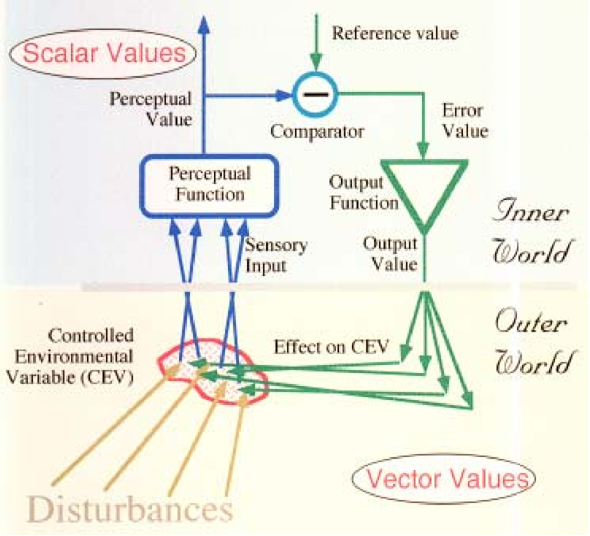


Figure H-1: An Elementary Control Unit. In a complex control system as envisaged in PCT, the perceptual signals of many control units form the inputs to the next higher level perceptual inputs in a perceptron-like hierarchy, while the outputs contribute to reference values at lower levels of the hierarchy or to effectors that act directly on the exterior world.

The ECU does one thing only. It accepts, from outside itself, data of indeterminate complexity and temporal extent that a "perceptual function" turns into a scalar value. This scalar value is called the "perceptual signal value". The ECU also accepts from outside itself a scalar "reference value" for its perceptual signal value. The perceptual signal value represents a state of the world outside the ECU, and the reference value constitutes the purpose or goal of the ECU, the value toward which the output of the ECU will tend to influence the outside world if control is effective. The difference between the perceptual value and the reference value is the "error value", which drives the output that affects the outer world, including the inputs from which the perceptual value is derived. If the output is such as to reduce the deviation between the reference and perceptual values, the whole circuit constitutes a control system in the engineering sense, no matter how complex or extended in time is the data that contribute to the perceptual value.

In the Powers theory, a complete complex control system is composed of many such elementary control units, arranged in a hierarchy of levels. At level N the inputs to any single elementary control unit are perceptual signals from elementary control units of level N-1, and the outputs from that ECU combine to form reference levels for elementary control units of level N-1. At level 1, the inputs are from the sensory systems, and the outputs are the effectors that act on the outer world. A higher-level control system thus has its effects in the world by setting the goals for several elementary control units at the next lower level. These try to bring the state of the world to match their new reference values; they do so by in their turn setting reference levels for yet lower-level control systems, until level 1 is reached, at which the inputs are from the sensory systems and the outputs are to muscles and other effectors.

H.2 RELATION TO OTHER FRAMEWORKS AND APPROACHES

H.2.1 Model-View-Controller

The VisTG Reference Model in its outline form (Figure H-2) suggests a three-level PCT structure. Each level represented by a grey loop is actually a complex perceptual control structure, though for most purposes this complexity can be ignored. To use the model, one need usually consider only that the visualising system in the human acts on some engines in the computer and receives displays created by other engines. In other words, although the human's objective is to understand some aspect of the dataspace, and perhaps to influence it, the actual work is done in the middle loop. This part of the VisTG Reference Model maps directly onto Trygve Reenskaug's original concept of the Model-View-Controller interface, shown in Figure H-3 [4][5].

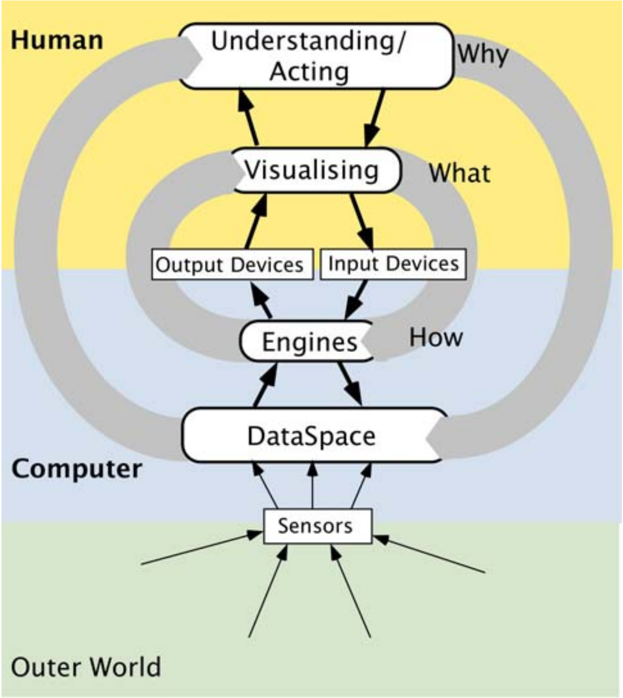


Figure H-2: The VisTG Reference Model in Outline Form, Showing the 3-Level Control Structure: The human wants to understand some aspect of the dataspace and possibly to influence it; the human accomplishes this (in part) by visualisation, using displays generated by engines taking data from the dataspace, controlled by other engines, some of which may also affect the content of the dataspace; the human actually interacts with the engines through physical interface devices.

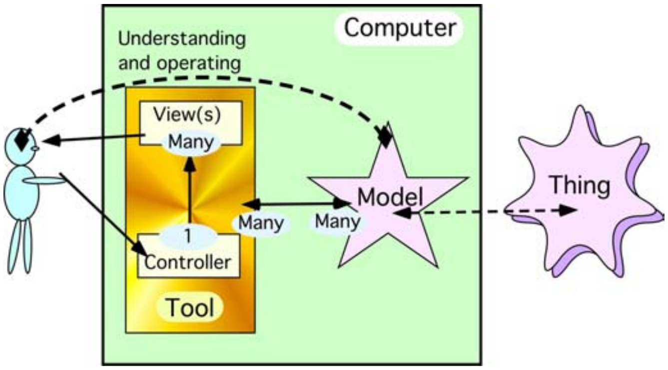


Figure H-3: Reenskaug's "Thing-Model-View-Controller" (based on diagram in http://heim.ifi.uio.no/~trygver/themes/mvc/mvc-index.html 30 August 2007).

The core of the VisTG Reference Model includes Reenskaug's Model-View-Controller view of a graphic user interface (GUI), even though the two approaches have quite independent antecedents. In the VisTG Model, MVC Controllers and Views are two kinds of Engine, and the MVC Model is the VisTG Dataspace. Sometimes, then, one can treat the VisTG Reference Model as though it were simply a Model-View-Controller framework, even though the VisTG Reference Model is more comprehensive, including as it does, engines for other purposes, such as to animate alerting daemons.

In the Final Report of IST-021 [6], the place of the Engines was described as follows:

Engines are of many different kinds, but there are two main classes:

- *Those that interact with data from the dataspace and manipulate them in some way, perhaps adding the results to the data space, perhaps displaying the results to the user, and*
- *Those that do not interact directly with the data, but work with the user in determining how the data should be selected or manipulated.*

> *Because the user needs to control the actions of the Engines, each individual Engine must be involved in its own feedback loop with the user. […] The user must be able to understand what an Engine can do and is doing, which may involve the user analysing or visualising the Engine's behaviour. The user must be able to instruct each Engine, and the Engine must be able to display to the user the necessary information that permits the user to determine how those instructions actually affect the actions of the Engine. Conventionally, to keep the picture simple, these individual user-to-Engine loops are omitted from diagrams of the VisTG Reference Model, but they must be considered when using the Model for system design or evaluation.*

The two main classes of Engine thus described are not directly equivalent to the MVC Controller and generator of a View, though they have similar functions. The first acts directly on the Model and perhaps provides the data for a View or possibly even generates the View; the second is a View Controller. Together they provide the functionality of the MVC Views and Controllers. Figure H-4 illustrates the mapping between the MVC concept and the VisTG Reference Model.

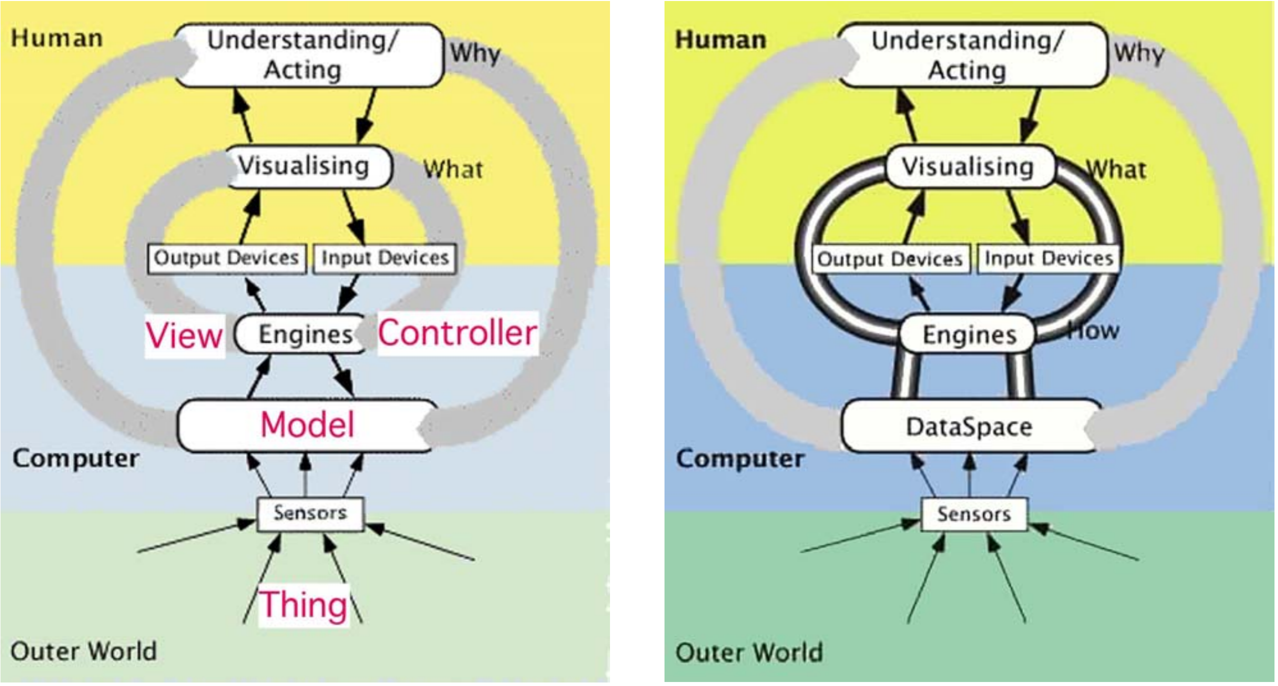


Figure H-4: Mapping the MVC Concepts onto the VisTG Reference Model – (a, left) Showing the mapping using the names of the MVC functions; (b, right) the VisTG Reference Model emphasising the core loop for most visualisations (reproduced from [1], Figure 4.1).

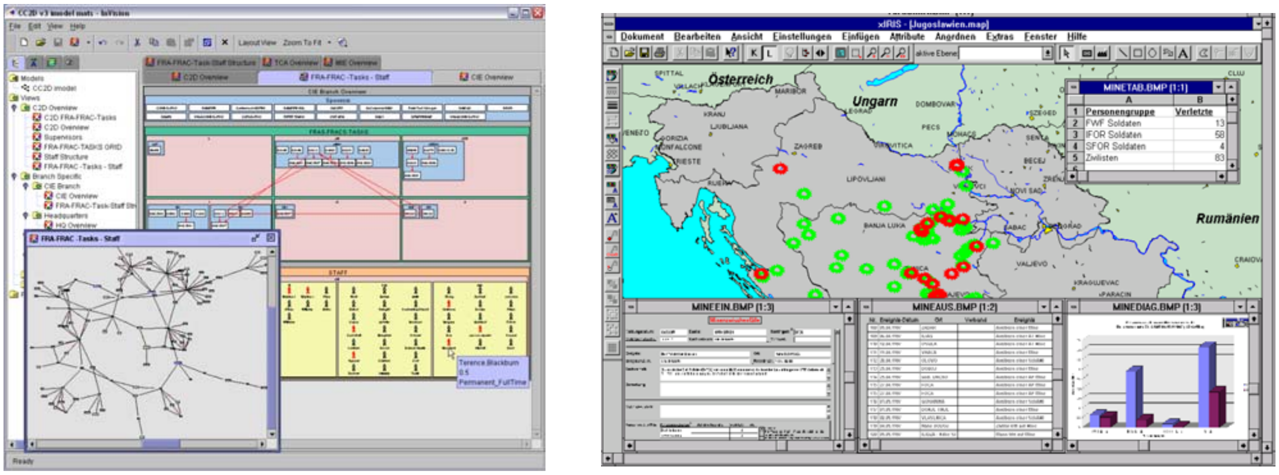
MVC often ignores the fact that the Model usually represents and is an abstraction of something in the outer world, although Reenskaug noted that fact in his 1979 concept of "Thing-Model-View-Editor" [4] (*Editor* soon became *Controller*, and *Thing* was quietly dropped). In the VisTG Reference Model, the relation between the data space and the outer world to which it refers is quite explicit; indeed the extended version of the VisTG Model includes a feedback loop through the real-world environment ([1], Figure 1-2).

The simple sketch of the VisTG Reference Model in Chapter 1 omits the outer world, just as does MVC when "Thing" is omitted. To do so may simplify some of the issues, by limiting the problems to those between the human and the computer, but that simplification may sometimes come at the cost of failing to address some real-world complication. Nevertheless, in much of what follows, we use this simplified form, both of the MVC structure and of the VisTG Reference Model. When we address the RM-Vis framework, however, the outer-world context does come into play.

As noted in Chapter 2, four modes of perception can be defined: Monitoring/Controlling, Searching, Exploring, and Alerting. The diagrammatic outline of the VisTG Reference Model seems to suggest that it is useful primarily for Controlling; this is far from the case. Returning to the historical antecedent of PCT, reference values can be such things as "I want to see Vienna", which would result in actions that affected a navigation Engine (a "Controller" in the MVC approach) to change the data selection and the display to include Vienna and display it in whatever manner had been selected previously.

The reason the user wants to see Vienna could be to see ongoing changes in the network surrounding or in Vienna (Monitoring/Controlling), to look for some aspect of Vienna need in support of some ongoing Controlling or Monitoring (Search), or to visualise something about Vienna for later reference when it might be needed in support of some future Controlling or Monitoring activity. These different reasons for wanting to see Vienna are likely to be best served by different kinds of display. If the Engines are sufficiently intelligent to create appropriate displays knowing the current perceptual mode, then the user must communicate the purpose to the Engine; if not, the user must use Controller Engines directly to affect the actions of the View Engines.

Many current display systems provide multiple simultaneous Views on the same Model (e.g. [7][8]). The multiple Views can be of many different kinds, as suggested by the examples in Figure H-5. They need not even be Views on the same Model, as suggested by Reenskaug's many-to-many link (Figure H-3) between the Model and the View-Controller complex he calls a "Tool". In the VisTG Reference Model, some Views may serve to control or monitor some changing situation, while others allow the user to explore the structure of the same or a different space for later reference. Each implies the existence of a control loop that depends on a user requirement. That requirement is of certain information to be displayed, which, in its turn, suggests a type of display. To create the required display is the job of an Engine, whereas to control it is the job of a different Engine.



MINETAB.BMP (1:1)

	A	B
1	Personengruppe	Verletzte
2	FWF Soldaten	13
3	IFOR Soldaten	58
4	SFOR Soldaten	4
5	Zivilisten	83

Figure H-5: Examples of Displays with Multiple Views on the Same Model – (left) from [7]; (right) from [8].

When there are simultaneous Views, each implies a separate loop Visualisation-Engine-Display-Visualisation. The VisTG Reference Model suggests to the display designer (who may be the user) that mutual interactions (beneficial or harmful) among the different views should be considered. Smestad [9] suggests an approach based on information theory to how those interactions might be analysed and translated into specific kinds of display. When we deal with the VisTG Reference Engine process, the interactions among views are considered explicitly (Section 1.3.2.2).

H.2.2 The RM-Vis Framework

RM-Vis, illustrated in Figure H-6, is described in detail in Annex G, including its specialization to network visualisation. In this section we consider only the way the RM-Vis framework integrates with the VisTG Reference Model as a stage in the development of the IST-059 Framework for network visualisation.

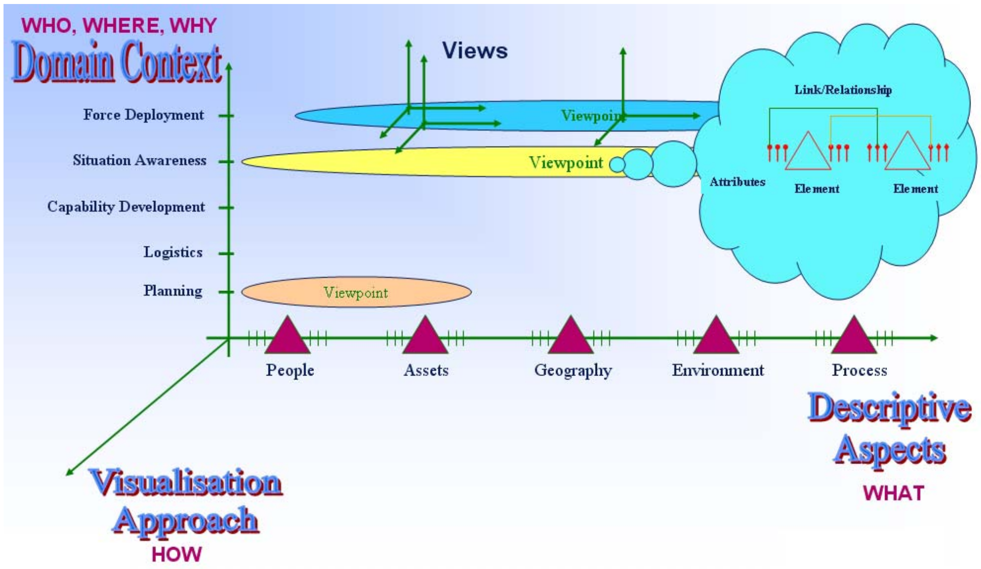


Figure H-6: The Three Main Axes of the RM-Vis Framework for Visualisation (repeated from Figure G-1 for easy reference).

Just as the MVC concept is concentrated on the computer side of the VisTG Reference Model, with a unitary human on the other side, the RM-Vis framework is concentrated largely on the human side, with a small element of the "Visualisation Approach" axis on the computer side. The "Domain Context" dimension of the RM-Vis framework deals with the social or military role of the user, which is the relationship of the user to the organizations in the outer world. That aspect is not a part of the VisTG Reference Model, except that it provides a set of possible answers to question Q1a of the canonical list of questions in the VisTG Reference Model process (from [6]):

- Q1. What user purpose is being considered?
  - Q1a. What higher-level purpose does this one support?
- Q2. What information does the user need to get from the computer to achieve the purpose?
- Q3. What does the user need to tell the computer to allow it to provide the needed information?
- Q4. What impediments might inhibit the user from taking advantage of the information provided?
- Q5. What impediments might inhibit the user from providing the computer the information it needs?
- Q6. Is there any mechanism to alert the user to information that might be important for the purpose but that is not currently evident in the display?

The second axis of the RM-Vis Framework is labelled "Descriptive Aspects". This axis deals with what the user wants to see. Choices on this axis provide possible answers to Q1 of the canonical questions, and when put together with a "Viewpoint" defined by positions on both the "Domain Context" and the "Descriptive Aspects", it also suggests answers to Q2. These first two axes then point the way to a domain- and task-dependent taxonomy of answers to the first two questions of the VisTG Reference Model process.

The third axis of the RM-Vis Framework is more complex, as shown in Figure H-7. It deals with the possibilities for presentation, which would correspond to the "Views" in the MVC approach, or to the Engines in the VisTG Reference Model, though at a lower level than is addressed in Q1 and Q2. It deals with the "How" of presenting the data. The axes, as shown in Figure H-6, sketch four dimensions that deserve consideration, with a few possible answers noted on each. These do not correspond directly to any element of the VisTG Reference Model but would be applied in the implementation of the part of the VisTG Reference Model's middle loop that leads from Engines to Visualisation.

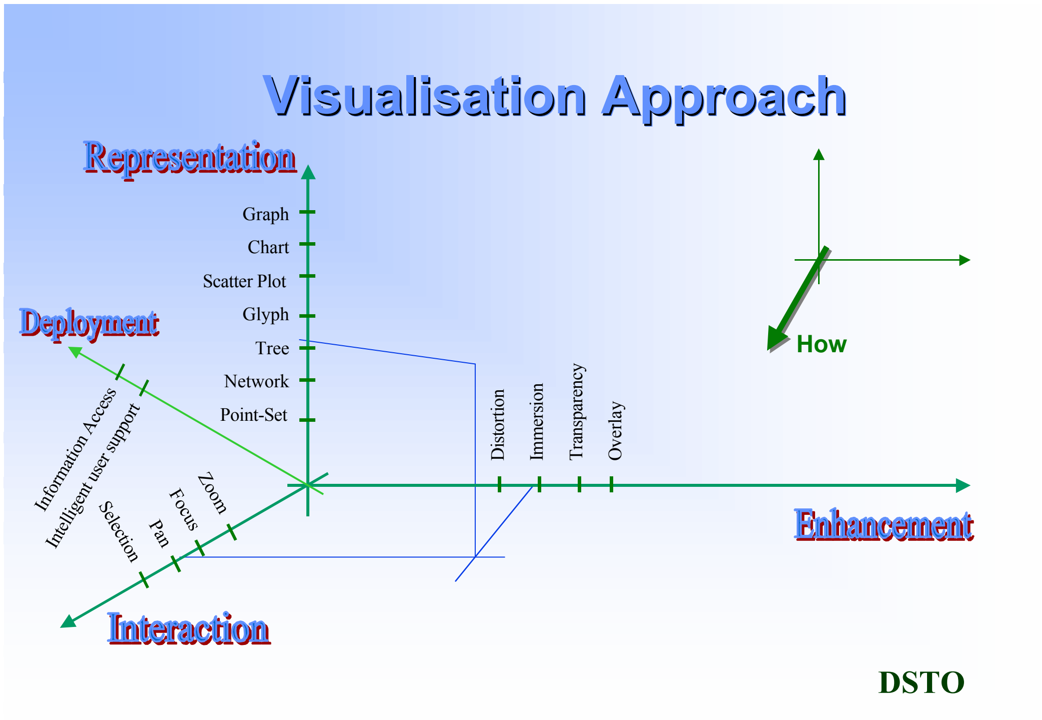


Figure H-7: Expansion of the "Visualisation Approach" Axis of the RM-Vis Framework (Slide 14 of Vernik's PowerPoint presentation in [10], republished in [1], Figure 7.2).

H.3 THE VISTG REFERENCE MODEL AND THE IST-059 FRAMEWORK

The RM-Vis Framework offers the start to a taxonomy for answers to the first two of the six questions in the VisTG Reference Model process. MVC replicates one level of structure on the computer side of the VisTG Reference Model itself. Neither addresses any of the last four questions, all of which are important. The IST-059 Framework for Network Visualisation incorporates the VisTG Reference Model and the RM-Vis Framework. The remaining questions must therefore be considered.

H.3.1 Internal Hierarchic Structure of the VisTG Reference Model

Before addressing the remaining questions of the VisTG Reference Model within the Framework for visualising networks, its internal structure must be sketched. The structure is described in detail in [6]; the present section should serve merely as a reminder and a pointer to the full description.

Remember that the VisTG Reference Model is based on W.T. Powers' Perceptual Control Theory [1]. Actions, at any level that is not connected directly to physical effectors (muscles, in the human, displays in the computer) are executed by setting reference values for lower level control systems (Figure H-8). Within the VisTG structure, the active elements are the Engines (and in the human the Visualisation level functions). The human senses no entities or structures directly. They are visualised by construction from the various sensations delivered by the eyes, ears, and other sensor systems.

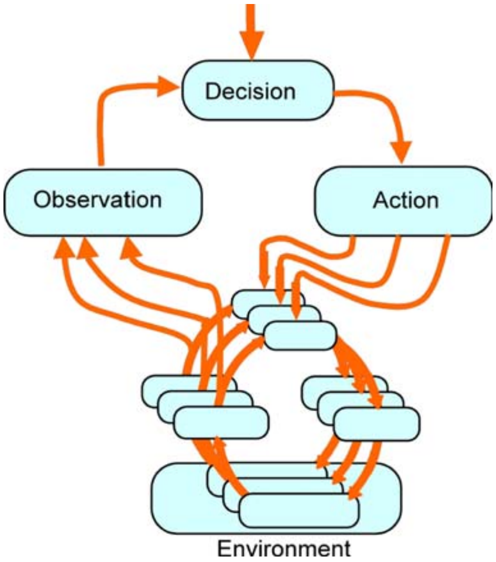


Figure H-8: Multiple Purposes Supporting One Higher-Level Purpose. The higher-level purpose is symbolized by the topmost downward-pointing arrow, and the supporting purposes by the three arrows leading from the higher-level Action. (From [6] Chapter 2, Figure 2)

However, as Figure H-9 suggests, what the human visualises corresponds more or less accurately to structures in the dataspace (the same happens when perceiving complex structures and events in the real world). When the person acts, usually the feeling is not of a series of muscle tensions, but of acting directly on the structures and entities in the world (or, if the displays and interaction mechanisms are well made, in the dataspace). These are shown as the "Virtual Connections" that form the higher-level loops in Figure H-8. The lowest level in Figure H-9 corresponds to the I/O, innermost, loop in the VisTG Reference Model, the highest level corresponds roughly to the structures that the person wants to visualise using the engines, and the middle level corresponds roughly to the engine loop itself, which controls what data are presented and how the presentation is done. Figure H-8 and Figure H-9 together may suffice to illustrate the hierarchic nature of the internal structure hidden within the thick grey arrows of the usual VisTG Reference Model diagram (Figure H-2).

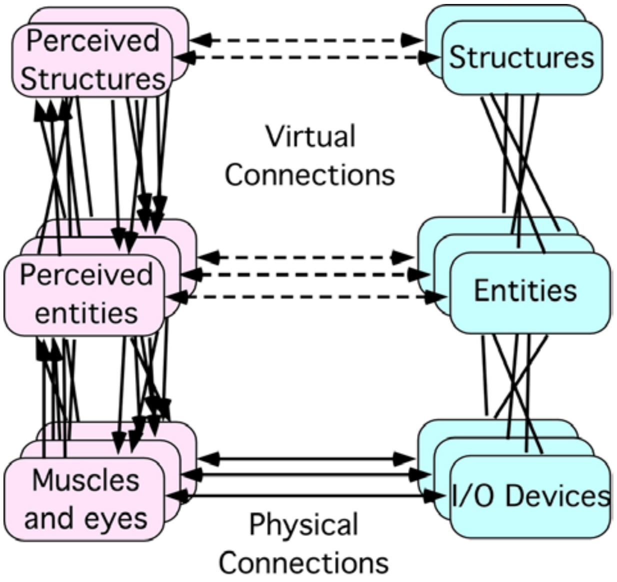


Figure H-9: The Human Visualises and Acts on Structures in the Data Space (or in the Real World), by Perceiving and Acting on Entities, and Perceives and Acts on Entities through Sensor Systems and Muscles that Interact Physically with I/O Devices (or with Real-World Objects).

The RM-Vis Framework addresses the first two questions (three, if Q1a is considered a separate question) of the six that are associated with each loop at each level of the VisTG Reference Model:

- Q1. What user purpose is being considered?
  - Q1a. What higher-level purpose does this one support?
- Q2. What information does the user need to get from the computer to achieve the purpose?

The four questions not addressed by the RM-Vis Framework are:

- Q3. What does the user need to tell the computer to allow it to provide the needed information?
- Q4. What impediments might inhibit the user from taking advantage of the information provided?
- Q5. What impediments might inhibit the user from providing the computer the information it needs?
- Q6. Is there any mechanism to alert the user to information that might be important for the purpose but that is not currently evident in the display?

All six questions should be addressed in some way in a complete framework.

H.3.2 The VisTG Reference Model and the IST-059 Framework

The IST-059 Framework is for visualising networks. The VisTG Reference Model is not so specialized, but it provides the frame for the Framework. In what follows, the discussion is largely restricted to network visualisation. Most of the detail, however, is in Annex B, where different kinds of network and possible useful views on them are discussed.

Q3. What does the user need to tell the computer to allow it to provide the needed information?

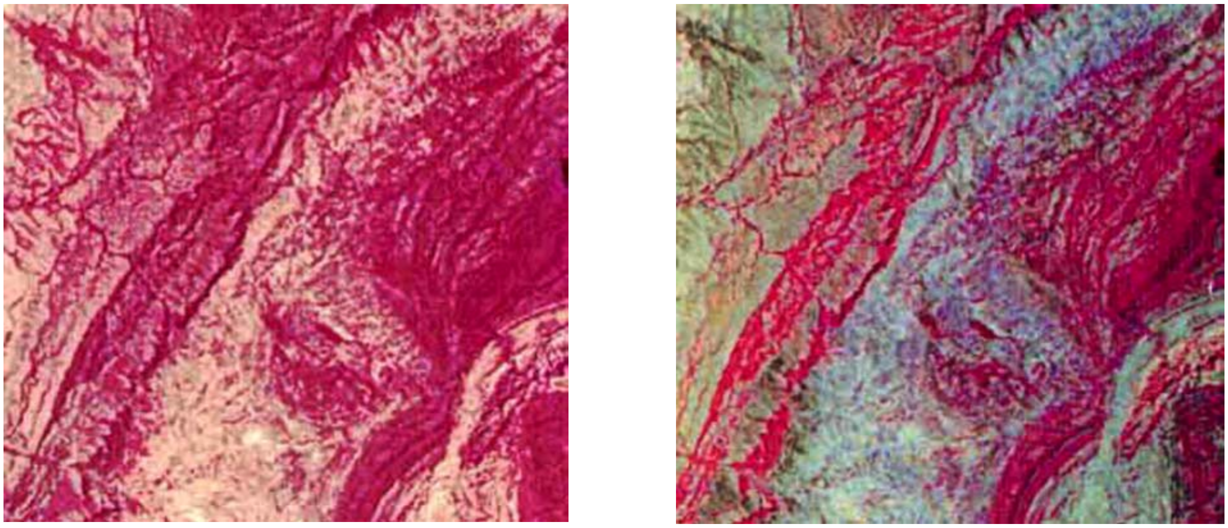
Question 3 is, in MVC terms, about how to address the Controller of views on a particular part of the network.

In different tasks and for different kinds of network, it might be easier to consider the MVC Model being Viewed to be either the whole network, or the part or properties of interest. Whichever the case, are at least two levels of control are to be considered. In the language of [6], a selection Engine needs to be told how to select the part of the network or the desired attributes for display, and a different Engine needs to know where and how to display: is the display for on-line monitoring, for Searching (in which case the selection engine will presumably be controlled interactively to allow the user to examine different parts or different attributes of the network), for Exploration (which also probably implies interactive control of the selection Engine), or for Alerting (in which case the display Engine needs to be given adequate information to allow it or a hidden daemon to identify when and where in the network the alerting condition occurs, and to allow it to know what other displays are ongoing when an alert is to be shown; this latter will affect the manner of the alert).

Q4. What impediments might inhibit the user from taking advantage of the information provided?

Question 4 represents a warning to the designer and to the user. A common impediment in displays is *masking*. Something displayed makes it harder for the user to take advantage of something else that is displayed. The masking may be at any level of analysis, from the simple clutter that can happen when a ball-and-stick display of a network tries to show too large a network on too small a screen, to the misleading of attention that happens when unwanted information is shown clearly, drawing the user's attention away from what should be the focus of the visualisation. The warning is to be careful when creating displays in which the user is expected to focus on local attributes rather than on the global pattern of what is being shown.

There are, of course, impediments other than masking. The display may contain the information desired, but contain it in a way not accessible to the user. Bjørke (Annex D) has shown how to analyze displays so as to maximize the possibility of information transmission when like elements are placed in the display, and has used it to generate map displays at dynamically varying scales. Taylor (Figure H-10, from [10], reproduced in [1]) has shown the use of colour variation to maximize information transmission. In Figure H-10, the information content of the two images is technically almost identical, but the colour variation in the left panel is not matched to the sensitivity of the human visual information transmission channels, whereas in the right panel the match is near optimal. In the left panel of the figure, the terrain in the top-left and bottom-right looks the same as the strip from the top-right to the middle-bottom, whereas these three areas are all distinctly different in the right panel. The data are the same in the two panels, the only difference being that the right panel uses a colour coding based on information theory.



a) b)

Figure H-10 (Reproduced from [1] Figure 2.3 and repeated from Figure B-8): A Multispectral Satellite Image of an Area of the Canadian Arctic in Summer – (a) As normally displayed in “false colour,” using one sensor channel as red, one as green, and one as blue; (b) By displaying the first three principal components of the spectral variation as, respectively, brightness, red-green contrast, and blue-yellow contrast. Several terrain differences that are invisible in Figure H-10a, are evident in Figure H-10b, even though both images display essentially the same data. (Images produced in 1976 by M.M. Taylor, then at DCIEM, Toronto)

Impediments to the user’s ability to visualise the desired information from a display more mundane than masking and poor matching to the human sensory systems may exist. In a multiwindow display, one window may simply overlay the part of another window that carries the important information. This is a kind of masking, but the problem lies outside, rather than within the View in question. Hence, the question and its answer imply the need for some kind of coordination among the Engines controlling of different Views, even if they are Views on the same MVC Model.

Q5. What impediments might inhibit the user from providing the computer the information it needs?

This question relates to the control loop through which the user communicates with the Engines, of whatever kind. In Figure H-8 and Figure H-9, it deals with the lower levels of the structure, where the user is not perceiving and influencing the implications of the data space, but is instead perceiving and influencing the state of one or more Engines.

The perceptual mode most common when dealing with Engines is Controlling, with Monitoring close behind. This implies that some part of the display must be dedicated to showing the state of the Engine, even if only temporarily (such as by the use of a pop-up window that might contain a menu). In the case of a View Controller Engine, the user might perceive its state indirectly, through the View being controlled. The state of a selection Engine might also be discernable through the associated View, but often some other view onto the Engine is required. Any information displayed to the user about the state of as Engine detracts from the information available about the data space, quite apart from drawing the user’s attention away from the real task at hand. Accordingly, it is normal to minimize the information shown to the user about the Engine state, and that minimization might inhibit the user from being able to effectively control the Engine state.

A common kind of impediment inhibiting the user from providing the information the computer would need if it is to be able to create the most useful views onto the data space is a limitation on the input devices and their associated degrees of freedom. It is difficult, for example, to use a 2-D mouse to select a region in a 3-D display space. Pointing devices are not very good for choosing displayable attributes of a network, but the use of a keyboard to describe the attributes requires display real-estate, and is often cumbersome when a variety of different attributes is required.

If the human user is to define the kind of display, then probably the Engine will show a menu of options, and that menu is likely to be hierarchic, since the options for a pie-chart display are very different from the options suited to a ball-and-stick display, and those in turn very different from the options for a matrix display of a network. These menus take space and time, as well as drawing the user's attention away from the network of interest.

Some of these impediments can be reduced if the selections and views can be predefined, allowing the user to spend time configuring the displays and only then turn attention fully to the network task. Some tasks may be sufficiently stereotyped that the arrangement of displays and windows can be specified in advance; for novice users, this is often a good strategy. Whatever the case, for a particular problem, Q5 cannot be ignored.

Q6. Is there any mechanism to alert the user to information that might be important for the purpose but that is not currently evident in the display?

Question 6 seems to lie outside the control systems depicted in Figure H-8 and Figure H-9, and it is not implicit in either the MVC or the RM-Vis structures, both of which deal primarily with the Controlling/ Monitoring mode of perception, while allowing, by default, for Search and Exploring. Alerting concerns attention. An Alert may bring the user's attention to something already being displayed, or it may signal that some new display might prove useful for the task at hand. In network analysis, Alerts are likely to be useful in searching large structures looking for subtle patterns of interaction among components, whether they be structural or of traffic.

It is hard to be specific about alerting systems. The possibilities are endless as to what patterns or events a user may want to set as an alerting condition. There are not so many possibilities for how to present an alert to a human – auditorily, tactually, or visually – but there are enough to make pre-specification difficult. The main characteristic of a successful alerting display is that to attend to it and to decide whether to change the focus of one's attention takes very little effort, especially if the decision is that there is no need to change focus. Furthermore, that "effort" is cumulative, so if a type of alert happens often and the decision is usually not to change focus ("crying wolf"), the alert becomes not an aid, but an impediment to the user taking advantage of the information displayed (Question 3).

H.4 THE VISTG REFERENCE MODEL AND THE FRAMEWORK PROCEDURE

The Framework procedure outlined in Chapter 2 requires the user to answer a series of questions that are designed to clarify the actual task requirements, and, in a fully implemented Framework, to ease the generation of queries to the Survey database of available software. The VisTG Reference Model suggests a series of useful questions. They can be simply summarized as "Why? What's needed? What's missing from what's needed? How to find it? What problems? What Alerts?"

Of course, in many situations, the user will be unable to answer all the questions in a way that can translate easily into database queries. Nevertheless, the attempt to answer them should clarify for the user exactly what needs to be done. Failure to answer "What's missing" may suggest a need for Exploration of the network structure, whereas a good answer to "What's needed" might argue for an application that supports Search, and has the ability to provide Alerts when the needed patterns are detected.

Whereas the VisTG Reference Model is applicable to any computer-based visualisation problem, the IST-059 Framework refers specifically to visualisation that involves networks. Network issues are considered throughout this report, but especially in Chapters 2, 3, 4, and 5, and in Annexes B, C, D, and E.

H.5 REFERENCES

[1] NATO IST-013/RTG-002, (2001), "Visualisation of Massive Military Datasets: Human Factors, Applications, and Technologies", *Technical Report, RTO-TR-030*, NATO Research and Technology Organisation.

[2] Powers, W.T., "Behavior: The Control of Perception", Aldyne, Chicago, 1973.

[3] Taylor, M.M., "Editorial: Perceptual Control Theory and its Applications", *Int. J. Human-Computer Studies*, 1998, 50, 433-444.

[4] Reenskaug, T., "Thing, Model, View, Editor: An Example from a Planning System", *Xerox PARC Technical Note*, 12 May 1979.

[5] Reenskaug, T., "Models, Views, Controllers", *Xerox PARC Technical Note*, 10 December 1979.

[6] NATO IST-021/RTG-007, (2007), "Multimedia Visualisation of Massive Military Datasets", *Technical Report, RTO-TR-IST-021*, NATO Research and Technology Organisation.

[7] Vernik, R., Von Itzstein, G.S. and Bouchard, A., "Imago: An Integrated Prototyping, Evaluation and Transitioning Environment for Information Visualisation", pp. 17-22, *11th International Conference Information Visualization (IV '07)*, Zurich, Switzerland, July 04-July 06, ISBN: 0-7695-2900-3.

[8] Kaster, A. and Kaster, J., "Visualisation of Heterogeneous Military Data in Geographical Information Systems: Architectural Concepts for Flexible User Interfaces", In *Proceedings of Workshop Visualisation of Massive Military Datasets*, NATO, RTO-MP-050, 2002.

[9] Smestad, T., "Outline for Concepts and Principles of Effective Visualisation", *Annex A to RTO-TR-IST-021 – Final Report of IST-021/RTG-007 Multimedia Visualisation of Massive Military Datasets*, 2007.

[10] Taylor, M.M., (1972), "Perceptual Principles Related to Remote Sensing", In *First Canadian Symposium on Remote Sensing,* Ottawa, 497-503, (Also DCIEM Research Paper 828).

Annex I – SOCIAL NETWORK ANALYSIS AND SURVEY TAXONOMY GLOSSARY

This glossary contains terms used in Annex E, Social Network Analysis, and terms used in Annex F, Survey Taxonomy and Analysis. It is not intended to be a complete list of terms, but contains many of those found in the product survey, and as such may help the reader in understanding a product's capabilities.

I.1 FUNDAMENTAL TERMS

Edge – The graph-theoretic term for link, connection or tie in a graph or network.

Entity Class – The type of items we care about, e.g. the set of actors.

Entities – The members of an entity class, e.g. individual actors.

Link – A specific relation among two nodes (also referred to as a connection or tie).

Network – A set of links among nodes such that nodes may be drawn from one or more entity classes and links may be of one or more relation classes.

Node – A specific entity, e.g. Joe, Martha.

Relation – A pairwise association, R, among entities a and b, e.g. by one entity being a parent of another, used to link nodes representing these entities into a network of individual relationships (links).

Relationship – An instance, denoted aRb or Rab, of a relation, R, among entities a and b linking a pair of nodes representing these entities into a network.

Vertex – The graph-theoretic synonym for node in a graph or network.

I.2 NETWORK STRUCTURE TERMS

2-Mode – A 2-mode network contains nodes that are of two distinct entity classes. For example, a social network's 2-mode network may contain actor nodes and event nodes.

Adjacency Matrix – A two-dimensional matrix definition of a semantic (e.g. social) network indexed in each dimension by the nodes of the network. Each entry indicates the presence (or quantifies the strength) of the association between the nodes indexing it.

Alternating Network – A network in which there are sets of nodes of different classes, such that no link can connect nodes of one class to other nodes of the same class, for at least one class of node.

Cellular Network – A topology featuring cliques connected by cell "leaders" and characterized by many cycles.

Complete – Complete graphs or networks have an edge (a link) between any two vertices (nodes).

Connected – Connected graphs or networks have paths between any two nodes, so complete graphs are connected.

Core-Periphery Network – A topology with many low-degree and few high-degree nodes linked into cliques featuring one large cluster and high centralization.

Discrete Graph – A graph with no edges (links), just isolated vertices (nodes); also called a null graph.

Hamiltonicity – A network has the property of Hamiltonicity if one or more topological cycles/circles (called a Hamiltonian) contains all of its nodes.

Hierarchy Network – A topology with no cycles among nodes, an in-degree of at least 2 and high centralization.

Matrix Network – A topology with no cycles among nodes, multi-in-degree and high centralization.

Multimode Network – A network where the nodes are in 2 or more entity classes.

Multiplex Network – A network where the links are from 2 or more relation classes.

Null Graph – A graph with no edges (links), just isolated vertices (nodes), also called a discrete graph.

Planar – Graphs or networks can be drawn in two dimensions without intersecting edges, a fact which simplifies their visualisation.

Random Network – A topology in which there is a normal distribution of degree.

Scale-Free Network – A topology in which there is a 1 / N (power law) distribution of degree.

Small-World Network – A topology in which many low-degree and few very high-degree nodes occur and in which there is no significant clustering.

Sink – A node that has 0 out-degree.

Source – A node that has 0 in-degree.

Symmetric Network – A network with an adjacency matrix, R, having the property that Ri,j = Rj,i, for node indices i and j.

Transitive Network – A network, G, wherein for any of its nodes a, b, and c, if there are links from a to b and from b to c in G, then there is also a link from a to c in G.

Tree – A connected graph with no closed paths.

I.3 NETWORK TRAVERSAL AND PATH-FINDING

All Pairs Shortest Path – The all-pairs shortest path gives the set of shortest paths between every pair of nodes (i,j) in the graph.

Breadth-First Search – A hierarchical graph traversal algorithm; starting at the root node, each node at the next level is visited until the desired end node is reached.

Circle/Cycle – A sequence of nodes in a graph or network formed by a closed path whose only repeating nodes are its first and last.

Closed Path – One that starts and stops with the same node.

Depth-First Search – A hierarchical graph traversal algorithm; starting at the root node, each node at the next level is visited until the desired end node is reached.

Diameter – The longest geodesic path in a graph or network between any two nodes.

Eulerian Path – An Eulerian path is a path in a graph which visits each edge exactly once.

Hamiltonian Path – A Hamiltonian path is a path in an undirected graph which visits each vertex exactly once. A Hamiltonian cycle is a cycle in an undirected graph which visits each vertex exactly once and also returns to the starting vertex.

Path – A walk in which no node is revisited.

Shortest Path – The single-source shortest path for node i is the set of shortest paths from i to all other nodes in the graph.

Topological Sort – Topological sorting orders the vertices and edges of a directed acyclic graph in a simple and consistent way and hence plays the same role for directed acyclic graphs that depth-first search does for general graphs.

Trail – A walk in which no tie (connection, link, edge) is followed more than once.

Walk – A traversal of multiple edges (links) in a graph (network), its length being the number of steps traversed.

I.4 NETWORK MEASUREMENTS

Centrality Measures – Measures of the relative importance of a vertex within the graph. There are several measures of centrality, including:

> **Betweenness Centrality** is a measurement of the extent to which a node lies between all other pairs of nodes on their geodesic (shortest) paths. Vertices that occur on many shortest paths between other vertices have higher betweenness than those that do not.
>
> **Closeness Centrality** is the shortest path between a vertex and all other vertices reachable from it. It is measured by the inverse of the sum of distances from a node to all the other nodes, which is then normalized by multiplying it by (n-1). For a directed network, each of in-closeness centrality and out-closeness centrality is measured separately, depending on whether the distances ‘from’ or ‘to’ other nodes are considered.
>
> **Degree Centrality** gives the relative degree (in- or out-) of a node. In SNA, this may be applied to discern key actors such as heads of hierarchies.
>
> **Eigenvector Centrality**, as defined by Bonacich (1972), of a node is (recursively) proportional to the sum of eigenvector centralities of the nodes it is connected to. It assigns relative scores to all nodes in the network

based on the principle that connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes. It is computed by the principal eigenvector of the adjacency matrix. Google's PageRank is a variant of the Eigenvector centrality measure.

Flow Centrality expands on the notion of betweenness centrality by assuming that actors will use all pathways that connect them, with a likelihood proportional to the length of the pathways. For a given node, the measure indicates how involved the node is in all flows between all pairs of nodes.

Information Centrality is a centrality measure based on the "information" contained in all possible paths between pairs of points. [Stephenson, Karen and Marvin Zelen "Rethinking Centrality: Methods and Examples." in Social Networks 11. (North-Holland: Elsevier Science Publishers B.V., 1989) pp. 1-37.]

Link Centrality, or edge centrality, is a measurement of the extent to which a link lies between all other pair of nodes on their geodesic paths.

Load Centrality is the fraction of the number of shortest paths that go through each node [NetworkX].

Cluster Recognition – Cluster recognition is any method that allows for the partitioning of a data set into sub-sets (clusters), so that each sub-set contains data that is related by some common feature. Some terms related to clustering include:

Bi-Component – A bi-component (or bi-connected component) of a graph is a maximal non-separable sub-graph. There are at least two different paths between any two nodes in the bi-component. It results from removing a cutpoint (articulation node) or a bridge.

Clique – A clique is a graph in which every vertex is connected to every other vertex in the graph. Cliques in a network may overlap, i.e. a node can be member of more than one clique.

Cohesion – A local measurement of how well-connected a group of nodes is, e.g. how many nodes must be removed to result in disconnection of the network.

Cohesive Block – Hierarchical (Nested) cohesive sub-groups made by removing 'node cut sets' recursively. (Netminer)

Community – Produces a nested structure of community by recursively removing the link with the maximum betweenness value until no links remain, and applying hierarchical clustering. [Michelle Girvan and M.E.J. Newman, (2002), "Community structure in social and biological networks"]

Component – A Component is a maximal connected sub-graph of a graph. [NetMiner]

k-core – A k-Core is a sub-graph in which each node is adjacent to at least k other nodes in the sub-graph. That is, for all nodes in the sub-graph, the minimum nodal degree within the sub-graph is k.

k-plex – A k-plex is a maximal sub-graph in which each vertex of the induced sub-graph is connected to at least n-k other vertices, where n is the number of vertices in the induced sub-graph. [http://www.tuta.hut.fi/studies/Courses_and_schedules/Isib/TU-91.V/2006/session_3.pdf]

Lambda Set – A lambda set is a maximal sub-set of nodes who have more edge-independent paths connecting them to each other than to outsiders. [LS sets, lambda sets and other cohesive sub-sets. By S.P. Borgatti, M.G. Everett, P.R. Shirey, Social Networks 12 (1990) pp. 337-357]

Minimum Cutest – Minimal node sets that make two or more components when removing them.

n-Clan – An n-clan is an n-clique which has diameter less than or equal to n as an induced sub-graph. [http://www.tuta.hut.fi/studies/Courses_and_schedules/Isib/TU-91.V/2006/session_3.pdf]

n-Clique – An n-clique of an undirected graph is a maximal sub-graph in which every pair of vertices is connected by a path of length n or less. [http://www.tuta.hut.fi/studies/Courses_and_schedules/Isib/TU-91.V/2006/session_3.pdf]

Node Set – A collection of nodes that group together for some reason, e.g. by virtue of a shared attribute.

Connection Measures – Connection measurements give an indication of how well-connected the nodes are to one another. Some connection measures include:

Accessibility – A measure of nodal interdependency, giving the probability that information will be transmitted by at least one of the one-step or two-step paths connecting node i and node j [Friedkin91].

Connectivity – Node connectivity of node pair (i, j) is the minimum number of nodes that must be removed to completely disconnect i from j. Similarly, link connectivity is the minimum number of links that must be removed to completely disconnect i from j. Network link connectivity is the minimum link connectivity between any pair of nodes, i.e. the minimum number of links that must be removed to disconnect the network. [Netminer]

Degree – The total number of edges/nodes to which a given node is connected.

Density – The density of a binary network is simply the proportion of all possible ties that are actually present. [Hanneman]

Dependency – A measurement of the extent to which i is dependent on j when going to other nodes.

Distance – The length of shortest path between two nodes.

In-Degree – The total number of nodes that send an edge to a given node.

Link Connectivity – Link connectivity of a pair of nodes is the minimum number of links that must be removed to leave no path between two nodes. Network link connectivity is the minimum connectivity between any pair of nodes, i.e. minimum number of links that must be removed to make the network disconnected.

Maximum Flow – Maximum flow from a source node to a sink node is the maximum possible total flow utilizing all the paths, given the constraint of flow capacity for each link.

Minimum Spanning Tree – The minimum spanning tree (MST) of a graph defines the cheapest sub-set of edges that keeps the graph in one connected component. [http://www2.toki.or.id/book/AlgDesignManual/BOOK/BOOK4/NODE161.HTM]

Node Connectivity – The Node Connectivity of network is the minimum number of nodes whose removal results in a disconnected diagram. The Node Connectivity Matrix shows connectivity of each pair of nodes in the graph. That is, the minimum number of nodes whose removal results in a disconnection between two nodes.

Out-Degree – The total number of nodes that receive an edge from a given node.

Equivalence – Equivalence gives a measurement of the similarity of two nodes. For example, two nodes are structurally equivalent if they have the same relationships to all other nodes in the network [Borner]. Other equivalence measurements include regular equivalence and automorphic equivalence.

I.5 MATHEMATICAL TERMS

Bijection – A one-to-one function or mapping f : E → F of the set of entities E onto a set of entities F, i.e. for all and only members, $f \in F$, there are all and only members, $e \in E$ such that E e → f.

Crisp – An entity either is or is not a member of a class (as opposed to **Fuzzy**).

Epimorphism – A homomorphism that is a surjection.

Fuzzy – An entity has a membership of between zero and unity in a class. For example, a man of 190 cm may have a membership of, say, 0.8 in class "tall".

Homomorphism – A function or mapping, f : R → S, from one relation (network or graph), R, to another, S, satisfying the condition f(a R b) = f(a) S f(b) that f carries R into S or, equivalently, that f\preserves R as S.

Injection – A one-to-one function or mapping f : E → F of the set of entities E into a set of entities F, i.e. for member, $e \in E$, there is a member, $f \in F$ such that e → f.

Isomorphism – A homomorphism that is a bijection.

Monomorphism – A homomophism that is an injection.

Surjection – A function or mapping f : E → F of the set of entities E onto a set of entities F, i.e. for every member, $f \in F$, there is a member, $e \in E$ such that e → f.

Annex J – UNIFIED THEORY OF NETWORKS

M.M. Taylor

J.1 OUTLINE

The IST-059 Framework for Network Visualisation is based around a general model of user interaction with dynamic displays, called the VisTG Reference Model (Annex H). The VisTG Reference Model is not specialized to any particular application domain or to any particular kind of data, and to make it useful for any specific application domain, the nature of the domain and its data must be amenable to description. To this end, IST-059 developed a unified description of generalized real-world networks and their contexts.

The Unified Theory has three major themes:

- The internal properties of a network *sui generis*;
- The relations of the network with the real world in which it exists; and
- The relation of the network to the user wanting to understand something about it.

This note outlines some of the elements of each of these themes. All are further developed throughout the Final report of IST-059.

J.2 NETWORK PROPERTIES

J.2.1 Internal Properties

The internal properties of networks have been the subject of much work, under the headings of Graph Theory, Social Network Analysis, and the like. The dynamical properties of some kinds of network are studied under the heading of System Dynamics. All of these independently developed research areas are subsumed within the Unified Theory. Some are quite general, some deal only with particular kinds of networks or with particular attributes of networks, but all consider only networks as such, abstracted from the real-world context in which the networks exist.

The Unified Theory offers a set of descriptive dimensions for considering networks. The first dimension is the Local-Global dimension. Local properties are those of individual nodes or links, plus the interfaces to the in- and out-links of a node or of a link to the nodes at either end of it. More detail is in Annex B.

J.2.1.1 Local Properties – Links

The properties of a link also have several dimensions of description. The most important may be whether the link carries “traffic”. “Traffic” is anything that leaves one node and arrives at another, changing something about the recipient node without necessarily changing anything about the transmitting node. In other words, “traffic” is not necessarily conserved. Examples of traffic include the obvious, such as vehicles on a road or bacteria that propagate infection, and the less obvious, such as ideas passed from one person to another. Traffic may be all that defines the existence of a link in some cases, such as the passing of infection from one person to another, whereas the link may have a defined existence even if no traffic ever passes over it in other cases such as a road that nobody uses.

Links that carry traffic have properties such as transit lag, load capacity, usage, and so forth. They are responsible for the dynamical properties of a network.

Traffic-free links have no dynamical properties. They simply relate one node to another. An example of a traffic free link is "owns" in "John owns a car", where "John" and "car" are nodes in some conceptual structure. Another example is the hyperlink created by writing "*http://some.domain/somepage*" in a Web page.

Links may be simple (elementary) or compound (complex). A simple link cannot be sub-divided into parallel links that separately connect the nodes at its two ends. A compound link has internal structure; it consists of at least two simple links that connect the same two nodes. A compound link may be braided, in which case all its simple links are of the same kind, such as the lanes on a multilane highway, or distinct roads that all connect town A to town B. A compound link may be complex, in which case its simple links are of different kinds, some perhaps carrying traffic, others traffic-free.

J.2.1.2 Local Properties – Nodes

In a graph, a node is simply a place where links meet. This is the simplest form of node. In the real world, nodes can be as complex as any processor, transforming its input traffic into output traffic of entirely different kinds according to the states of other nodes connected by traffic-free links. To characterize nodes completely would be to characterize all software, and then to extend that characterization to all biology.

Although nodes can do almost anything with their incoming traffic, nevertheless some properties can usefully be described. In System Dynamics, nodes are "stocks", and have a value. In a Petrie Net, a node may not deliver an output until it has received some number of inputs. In general, if a node has both input and output traffic, there must be a temporal relationship between input and output such that output that depends on an input will happen at some later time than the input. Nodes with traffic-bearing links cause delay.

If all the links connected to a node are traffic-free, the node cannot be a processor. Until it is in some way related to the world outside the network, it is just a connecting point, perhaps labelled, but no more. Such connections occur through the semantic and pragmatic embedding fields described below (and see Annex B for more detail).

J.2.1.3 Global Properties

Global properties are properties of a network that cannot be ascribed to any single node or link, but are inherent in the way the nodes are interconnected in all or part of a network. For example, the diameter of the network is computed by finding the shortest path (succession of links) between a pair of nodes, doing this for all pairs of nodes, and then selecting the longest such minimum path. No one node has this property; it is a global property of the network.

Some global properties can be ascribed to individual nodes. These properties depend on the node's position in the network. For example, the number and proportion of minimum paths between node pairs that traverse a node can be computed for every node in the network. This proportion is one measure of the centrality of the node, but it is not an intrinsic property of the node, because it changes when the structure of other parts of the network changes.

Many global properties, both of the network as a whole, and dependent on the network but ascribable to individual nodes or links, are computed in mathematical graph theory, and in the discipline known as Social

Network Analysis (SNA). In graph theory, nodes are ordinarily considered only as the meeting places of simple links, whereas in SNA both nodes and links can have other properties and internal structure.

J.2.1.4 Network Types

Several different types of network are identified, in addition to the classical network in which well defined links interconnect well-defined nodes. Within the classical type of network are several sub-types, including "Striped". A Striped network is a sub-class of a multimodal network in which nodes not only have different classes, but there exists at least one class of node that cannot link to another of the same class, but must link to a member of some set of other classes of node.

In addition to the classic kind of network with point-to-point links, two kinds of "Broadcast" networks are defined, "Immediate" and "Stigmergic". Both are traffic-carrying and both are necessarily "Striped", having a class of node "Transmitter" that can link only a class of node "Receiver". In an Immediate Broadcast net, traffic that is not received at the moment it is available is lost forever, whereas in a Stigmergic Broadcast network, traffic remains available for reception for a substantial time period compared to the speed of propagation.

Networks may be homogeneous, all the links being of the same kind, or heterogeneous, having links of different kinds, such as point-to-point and broadcast.

Networks may have both crisp and fuzzy links and nodes. The fuzziness property is independent of any probabilistic considerations; it depends on how well the real-world entity fits the concept of "link" or "node". As an everyday example, a lightly travelled expressway is a link between two interchanges. If the expressway is blocked by a major accident, it is not a link. Between these two extremes, if the traffic is heavy and slow, it is neither clearly a link nor clearly not a link. It has a fuzzy membership less than unity and greater than zero in the class "link".

Networks differ in their topologies. The effects of topology are the province of Graph Theory, Social Network Analysis, and System Dynamics. The Unified Theory provides a place within its structure for these theories. One important topological consideration is whether a traffic-carrying network contains cycles, since cyclic networks can have much more complex dynamics than can acyclic networks.

J.3 REAL WORLD NETWORK PROPERTIES

Real-world networks can be described in two independent dimensions, their internal properties and relations as sketched above, and their relationships with the world in which they exist.

J.3.1 Syntax, Semantics, and Pragmatics as Linguistic Concepts

In Table J-1, the words "Syntactic", "Semantic", and "Pragmatic" are used by analogy with their meanings in linguistics. In linguistics, "syntactic" relationships among the words are those that concern generic word types such as noun, verb, and adjective, and involve relationships such as word sequence and phrase ordering. These relationships are entirely among the words. A word sequence such as "*Green furiously colourless sleep ideas*" is not syntactically well formed, but "*Colourless green ideas sleep furiously*" is. However, the latter sentence is not semantically well formed.

Table J-1: Two Dimensions of Network Description

	Syntactic (within the network)	**Semantic (supporting the network)**	**Pragmatic (meaning of the network)**
Local	Properties of individual nodes and links	Supporting environment of nodes and links	Interactions with the external environment of individual nodes and links
Global	Network or sub-nets	Supporting structures for the network or sub-net	Interactions of network or sub-nets with the external environment

Semantics is concerned with the usual meanings of words, and something cannot be at the same time colourless and green, ideas do not sleep, and "furiously" is not normally associated with how one sleeps. For a sentence to be semantically well formed, its words must be capable of conveying some concept, such as "*Pale green frogs sleep quietly*". It may or may not be true, but it is capable of being true, because green can be pale, and frogs can be that colour and perhaps can sleep quietly.

Pragmatics is concerned with the relation of the sense of the text to the facts of the real world. To say "Barack Obama had tea with President Roosevelt after the election" is well formed semantically, but is pragmatic nonsense. Without enquiring as to Barack Obama's true actions after the election, we know this must be false because of our background knowledge of the world. It simply does not fit, pragmatically, with what we believe could be true. Pragmatics is concerned with the way the world could be, knowing what we do about the way the world is.

J.3.1.1 Network Syntax, Semantics, and Pragmatics

One can talk about networks using an analogy to these rough approximations to the linguistic concepts of syntax, semantics, and pragmatics. Syntactic analysis and description of a network concerns only the relations among network entities within the network. A node is a node, whether it be labelled "John Jones" or "node 143". A link is a link, whether it represents a road, a family relationship, or a Web hyperlink.

The internal properties of a network may be local (pertaining to individual nodes and links) or global (pertaining to groups of nodes and links, which usually would be connected sub-nets of the whole network). Their relation to the external world may be nil (the domain of graph theory and much of Social Network Analysis – SNA).

Almost all network analysis is, in this structure, global and syntactic. Such analyses concern only the internal structure of the network. In Semantic Networks or Social Network Analyses, the nodes and links may be labelled with words that relate to real-world properties such as "Fido >isa> Dog". These labels have no effect on the analyses, but assist the human user's mind to make some sense of the network. Despite the labels, and despite the name "Semantic Network", the analyses are entirely of properties within the network.

Most real-world networks are supported by some substrate. The network of TCP/IP protocol connection among different computers is supported by the hardware of the computers and by the wires and wireless connections that convey the physical signals. The TCP/IP network itself supports the network over which the traffic is e-mail messages. The e-mail network supports many different social networks connecting humans, and so forth. The TCP/IP network supports other networks besides the e-mail network, one of which is the World Wide Web.

The support structures of a network constrain the behaviours of its nodes and links. The hardware limits the number of packets per second that can travel a particular TCP/IP link, and the physics limits how quickly a packet can get from one node to another. These limits constrain how fast a server can respond to a client request to be served a Web page. The analysis of these effects on the behaviour of the Web is essentially semantic. The content of the Web traffic is irrelevant, but the quantity and rate of the traffic must conform to the constraints imposed by the supporting structure. We take such constraints to be analogous to the constraints imposed on the interpretation of a sentence by the normal meanings of the words of a text.

In linguistics, “pragmatic” refers to the real-world context of the text. Likewise it is reasonable to treat the relationships of a network with the environment outside the network and its supporting structures as pragmatic. For example, the landscape over which a road network is constructed affects the nature of the road, and its value for different users. The landscape does not affect the nodes and links of the road network when it is considered as a graph. The network’s syntax is not affected by the landscape. The network’s semantics is influenced by the construction techniques of the road, whether it is a well-built highway or a casual cart track, but its pragmatics are affected by a wide range of environmental circumstances, from the social environment to the climate and weather. Which, if any, of these environmental contexts is important depends on the reasons the network is being studied.

J.3.2 Embedding Fields

The preceding discussion concerned matters strictly within the network (syntax), matters relating to the real-world supporting structures that constrain the real-world behaviour of the network (semantics), and the influence of the surrounding environment that might affect the network and be affected by the network in ways important for some study of the network (pragmatics). The latter two can be recast in terms of “embedding fields”, semantic and pragmatic.

J.3.3 Semantic Embedding Fields

The idea of an embedding field developed from a speculation that:

1) *A physical network always has the possibility that a conceptual network lies on top of it. The conceptual network may map homologously onto the physical network if the relationships between nodes are defined as such, but in most cases, the conceptual network involves only sub-sets of the physical network.*
2) *A conceptual network may exist without any underlying physical network.*

This speculation introduces a distinction between conceptual and physical networks. A conceptual network exists in someone’s mind, but it can conform to a physical network. If it does conform to a physical network, it may well map only onto a sub-net of the entire physical network. Following this speculation led to a more general appreciation of the relation between networks and their real-world support, which came to be understood as the semantic embedding field for the network.

It is possible not only for a conceptual network to be supported by a physical one, but also for a physical one to be supported by a conceptual one. Consider the following example:

> The network of possible airline connections derivable from published schedules is non-physical, but to implement a trip using the scheduled connections requires a physical network in which the nodes are airports and the links are defined by the traffic of physical aircraft. The network defined by the aircraft travel (traffic) is supported by the network defined in a published schedule, but with variations due to events not forecast in the schedule plan.

> There is a non-physical network in which the links between the airport nodes are defined by the actual trips taken by all passengers on a given day. If no passengers flew a particular scheduled link because of some statistical anomaly, this trip-based network does not include that link. The trip-based network has as its substrate a physical network whose links are defined by the actual aircraft flights between airports. This physical network itself depends on a non-physical conceptual network defined by the schedule plan. The difference between the trips planned on any given day (a conceptual network) and those actually taken (a physical network affected by delays and cancellations) indicates the importance of the intervening physical network. Here we have a case of a non-physical network that has a physical network substrate, which in turn is based on a non-physical network.

A supporting network constrains the supported network but does not define it. The two are often of entirely different character. Aircraft flights are not of the same type as links in a schedule. Yet the schedule constrains when and where aircraft fly. The constraint in this case is not absolute, but probabilistic. Aircraft can, but rarely do, fly where no flight is scheduled, or fail to fly where one is scheduled. For the most part flights occur where and when they are listed in the schedule, one flight occurs for each flight listed in the schedule, and that flight does not take off before the time listed.

In the airline example, the schedule is a semantic embedding field for the network defined by the flights that occur, and both that network and the conceptual schedule network are semantic embedding fields for the network of passenger trips that use air transport.

The hierarchic structure of semantic embedding fields has much in common with the hierarchic structure of inheritance in object-oriented programming (OOP). In OOP, a child object inherits all the properties of the parent other than those that are modified, and in addition can have properties of its own. In a traffic-carrying embedded network, the traffic flows are dictated or limited by the flows of the embedding network. Although passengers are very different from aircraft, nevertheless no passenger can fly unless an aircraft does. Although e-mail messages are of very different character from TCP/IP packets, no e-mail message can be transmitted unless several TCP/IP packets are passed. The structure of the network defined by e-mail messages is different from the structure of the network defined by TCP/IP packets, since any one message between A and B may be constructed of a set of packets that followed different multilink paths from A to B, but no link from A to B in the e-mail network can exist unless there is at least one corresponding path between A and B in the TCP/IP network.

The dynamic behaviour of a network is also constrained by any traffic-carrying semantic embedding fields, since a link cannot convey its traffic any faster than is permitted by the supporting network. Time delays are critical determinants of the dynamics of networks, particularly of cyclic networks; if the embedding network is acyclic, so must be any embedded network, at least in a single-inheritance hierarchy. Even in a multiple-inheritance hierarchy, in which a network is supported by two or more independent embedding networks (which would permit cycles when neither embedding field does), the timing delays of supporting links and paths are critical.

In traffic-carrying networks, both the traffic and the network structure are constrained by their embedding fields. In traffic-free networks, only the structure is constrained (or inherited). A network of “friendship”, for example, is constrained by an underlying network of “acquaintance”. There can be no friendship link between A and B if they are not acquainted. Accordingly, if there is an embedding hierarchy involving traffic-free networks, something can be learned about the embedded network by studying the embedding network.

J.3.4 Pragmatic Embedding Fields

A pragmatic embedding field includes any influences on the network from outside the network other than those that support the structure and activity of the network (the semantic embedding fields). For example, consider the traffic flows on a road network. One pragmatic embedding field is the larger social network among the people living in the area, many of whom go to work in the morning and go home in the late afternoon. This social phenomenon is not a property of the network, but the rhythmic variation in traffic density is. The social environment is thus a pragmatic embedding field for the road network. The effects of a 9 to 5 workday are nowhere to be found in the network or in any of its supporting structures. They are a pragmatic influence.

Another example is one that has concerned IST-059 and its predecessors – the effects of disruption of one infrastructure network (such as the electricity supply) on another (such as the water system or the food distribution system). These effects differ from those of the previous example, since they are strongest on a few specific nodes or links in the network of interest. Although the electricity supply supports food distribution, the electricity supply network does not support the food distribution network in the sense of being a semantic embedding field for it. The electricity supply network is part of the pragmatic environment of the food distribution network, and therefore a pragmatic embedding field.

Whereas semantic embedding fields are usually themselves networks, and thus of dimensionality 1.0, pragmatic embedding fields can have any dimensionality from zero up. A zero-dimensional embedding field could be the individual words used in a propaganda campaign to influence the network of popular opinion transferred by face-to-face or internet communication. A or high-dimensional embedding field might be the electromagnetic environment that affects a wireless communication network (three spatial dimensions, one time, and many electromagnetic frequencies, though only frequencies in the band of the intended wireless communication actually affect the transmission).

It is hard to characterize pragmatic embedding fields in the way we have learned to characterize the syntax, and in some cases the semantics, of networks. The two examples above illustrate the difficulty. There are, however, one or two useful dimensions of description:

- Does the pragmatic embedding field affect global network structure properties, local properties, or traffic properties?
- Is the influence of the pragmatic embedding field evenly distributed over the network or concentrated over a smaller sub-net?
- Is the influence of the pragmatic embedding field predictably time-variant (as in the rush-hour example), impulsive (as in a catastrophic failure), or unpredictably but continuously time-variant (as in weather effects on travel)?

When we treat the relationship of the network to the user, the pragmatic embedding field is taken to be the aspects of the network environment that affect the network in ways that influence the user's ability to understand the network in relation to the current task.

J.4 RELATION OF THE NETWORK TO THE USER

J.4.1 Entropy and Information

A network is, of itself, usually interesting only to an analyst. To most users, the network is no more than a milieu in which there is a problem. The problem is often one with a simple answer, though to find that answer may be very difficult. The network may be extremely complex, especially when considered in whatever pragmatic embedding field is important to the user's problem. In information-theoretic terms, it is of very high entropy, whereas the information the user wants is of very low entropy.

It is possible to follow the information-theoretic concepts at least qualitatively, and often quantitatively, through the chain from the real world to the user's mind, as suggested in Figure J-1. To transfer information from A to B is to introduce correlation between the two. A remains unchanged, but the structure of B is altered in some way such that something about A can be better inferred from B after the transfer than would have been possible before the transfer. The entropy of the combination of A and B is reduced when B is changed to better correlate with A, even though the entropy of B itself may be increased.

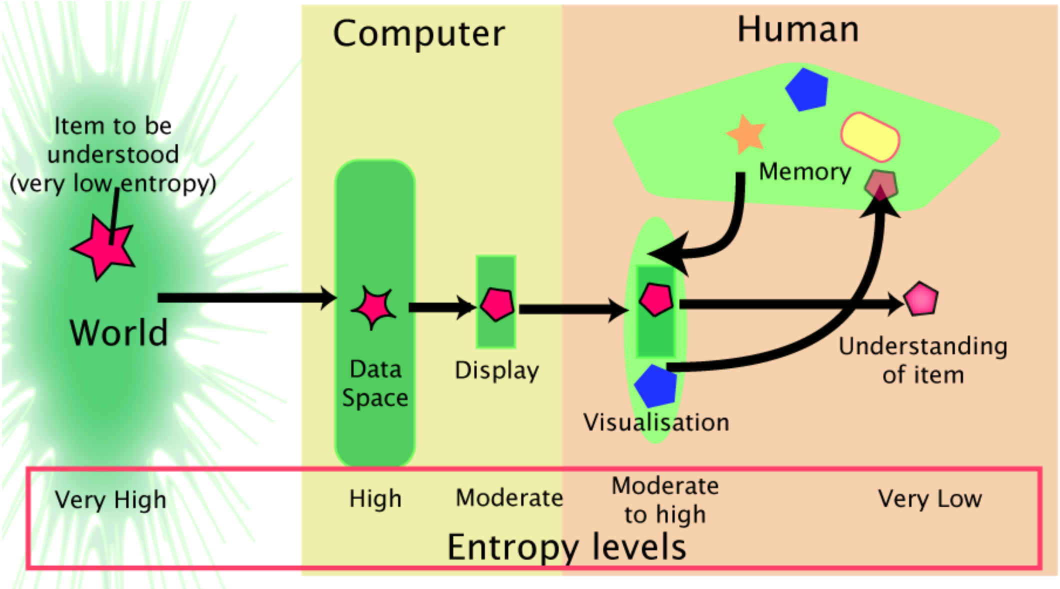


Figure J-1: Variable Entropy Levels at Different Stages of Processing from the Real World to the User's Understanding of a Low-Entropy Problem. Green areas represent entropy levels, red five-sided figures the low-entropy data relevant to the problem.

The real world is a very high entropy place, but only a relatively small amount of information enters the dataspace. That "small amount" nevertheless may generate a dataspace that is still of too high entropy for a display to show or for a user to take in at one time. Some of what is in the dataspace is relevant to the problem, but quite probably some relevant information in the world has not entered in the dataspace (as suggested by the different shapes of the red star in the two left sections of Figure J-1). The display usually can show only a small amount of what is available in the dataspace, and probably omits some of the data relevant to the problem, as suggested by smoothing out the star points in the Figure J-1.

When the human user looks at the display, the display content is combined with the user's background knowledge and experience to create a visualisation of relatively high entropy. This visualisation both depends on and feeds back to the user's memory, and, with luck, somewhere in it is a visualisation of how the problem fits in its context. This visualisation finally leads to the user's understanding of the problem, which is usually of very low entropy, requiring only the information implicit in the answer to a question such as "Who is the leader of that group?" or "Can my troops get there in time?"

Because in most cases the display cannot represent all that is in the dataspace, the user must interact with it to select different views on the dataspace, both by navigating the display through the space and by varying the algorithms used to extract and present information in a useful way. Interaction is a central component of the Unified Theory, and is schematically represented by a three-level structure of feedback loops known as the "VisTG Reference Model" (Figure J-2a and Annex H).

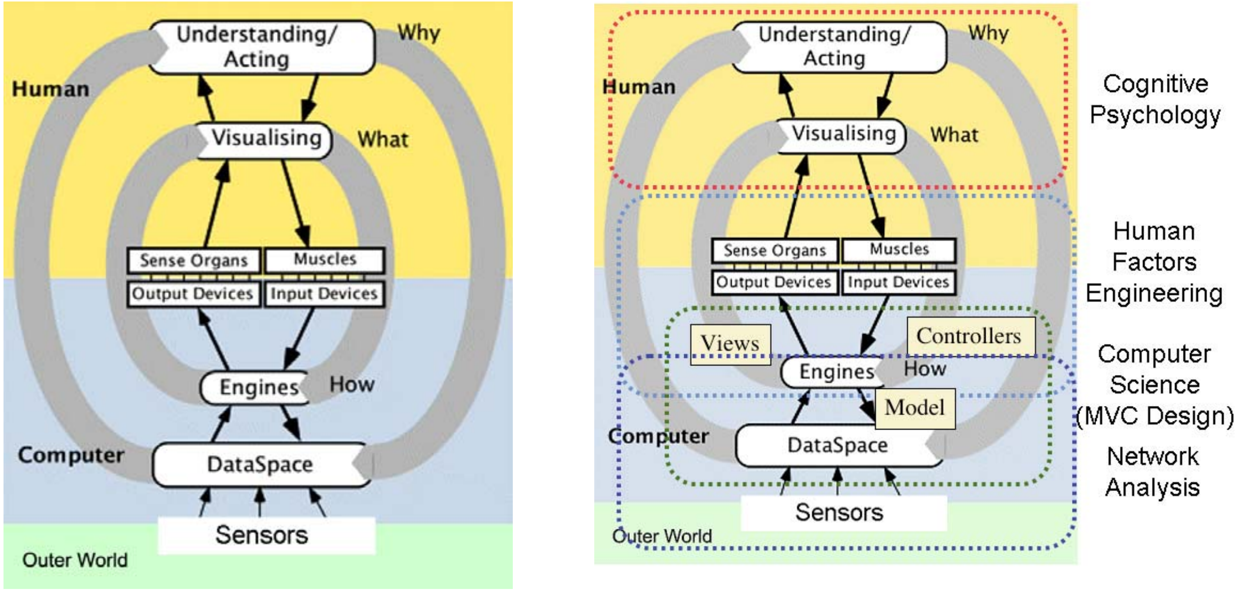


Figure J-2: The VisTG Reference Model – (a, left) showing the three nested feedback loops; (b, right) showing some of the disciplines valuable in instantiating different areas of the model.

The VisTG Reference Model has implications for several different disciplines when it comes to implementation, as shown in Figure J-2b. The Unified Theory places them in a coherent structure.

Figure J-2 represents the information flow of Figure J-1 in the series of upward pointing black arrows in the left middle of the figure. Figure J-1 omits the "Engines" stage, because Engines are processors that transform the data flow, and do not represent a stage for which an entropy measure is easily defined.

J.4.2 Modes of Perception

The entropy levels asserted for different stages of processing in Figure J-1 are not always true for a user's task. Sometimes a user wants to learn about the structure of the network in some detail. In such a case, the "problem" entropy can be almost as high as that of the whole dataspace. At other times, the problem entropy is small, because the user is trying to answer a simple question, or is tracking variation over time in some attribute of the network. These differences are incorporated in the Unified Theory by reference to four "Modes of Perception" (Annex B).

The four modes are:

1) **Monitoring or Controlling:** The person is attending to and perhaps influencing some aspect of the display in real time. The attribute being monitored may be overtly apparent or be derived from the user's visualisation or analysis, but it is a unitary thing. People are not good at tracking more than two

or three attributes simultaneously unless they vary slowly. Information bandwidth matters, but the static entropy of the set of monitored attributes is small.

2) **Searching:** When Monitoring, some data may be needed in order for the person to be able to determine the effective value of the attribute being monitored. The person must Search for the missing data, which is needed at that moment, in real time.

3) **Exploring:** The person is gathering information about the world that may be useful in some later Monitoring or Controlling. For example, exploring a network structure may enable an analyst to spot an anomaly very quickly when one occurs. Exploring is always done as a background activity, not as a real time interaction with the display or with the real world. The difference between Searching and Exploring may be illustrated by example. The person wants to write (a Controlling use of perception) but has no pencil. He Searches until he finds a pencil in a drawer, whereupon the Search stops. Another person wants to know what is in the drawers in case something useful might be there, and sees a pencil. The Exploration continues, but when she wants to write, she can pick up the pencil without Searching.

4) **Alerting:** Alerting in humans is a non-conscious background activity. An example is the quick eye-flick when one sees an unexpected motion in the visual periphery. The eye-flick is sufficient to allow the person to determine whether to shift attention to the place of the movement, or to something suggested by the fact of the movement. Many alerting filters are usually active at any moment. In computers, the same function can be performed by autonomous daemons that provide a signal to a human alerting system when the daemon discovers a condition that it has been programmed to detect, usually an event or a structure in a complex dataspace.

The modes of perception are important when determining the requirements for display. Most displays are constructed with Exploring as the main expected use. Displays suited for visualisation when Exploring can be quite complex, whereas displays for Monitoring/Controlling should be simpler, including only enough of the context of the Monitored attributes to help the user keep oriented within the dataspace.

J.4.3 Modes of Interaction

Four levels of interaction are identified in the Unified Theory. These are discussed in detail in Annex B. The VisTG Reference Model (Figure J-2) is most pertinent to the closest interaction, simply called “Interactive”. The four levels are:

1) **Interactive**: A single end-user has direct control of the display in real time.

2) **Coordinated**: Several users all have direct control of aspects of the display content, and must coordinate who controls what when.

3) **Mediated**: An operator controls the display on behalf on one or more end-users.

4) **Passive**: The end-users have no influence on the content of the display, as is the case for displays published in books or journals, presented in lectures, or in TV programming. A briefing might be mediated, but is more likely to be passive.

The interaction mode has implications for display. For example, an interactive or coordinated display may have a looser structural syntax than would be appropriate for a mediated or passive display. Interactive displays may usefully be more complicated than mediated or passive displays, if the interaction allows the user appropriate manipulations.

J.5 CONCLUSION

The Unified Theory offers a structured way to understand networks, especially in the real world beyond the abstract graphs of Graph Theory or the labelled graphs of Social Network Analysis. It incorporates syntax, semantics, and pragmatics of networks to augment existing mathematical and psychological approaches to network analysis and visualisation, and provides a structured place for those approaches.

Annex K – 2005 VISUALISATION NETWORK-OF-EXPERTS WORKSHOP

Visualisation Network-of-Experts

Supporting NATO Research Task Group IST-059 on Network Visualisation

http://www.visn-x.net

Cologne Cathedral

Workshop:

Social Network Analysis and Visualisation for Public Safety

18th – 19th October 2005

FGAN-Research Institute for Communication, Information Processing & Ergonomics (FKIE)
Neuenahrer Straße 20, D-53343 Wachtberg-Werthhoven, Germany

Edited by M. Varga

The NATO Research Task Group "Visualisation Technologies for Network Analysis" has invited the Visualisation Network of Experts (Vis N/X) to advise it on how social network analysis and visualisation can support public safety. Vis N/X has accepted this invitation.

CTRL, Click on the *blue*, *italicised* links to view the associated files.

Agenda

Tuesday, 18th October 2005

0915	Dr. Jurgen Grosche, FGAN	Welcome to FGAN
0930	Annette Kaster, FGAN	Information on local arrangements
0945	Round Robin	Introductions
1000	Vincent Taylor	RTG-025 and the N/X
1015	Zack Jacobson	*Workshop Introduction – Social Network Analysis for Public Safety*
1045	**Coffee Break**	
1100	**Provocations and Discussions**	
	Mark Nixon, Martin Taylor, Margaret Varga	*Network Visualization: Reference Model*
	Marcus Lem	*Developing Frameworks for Data Representation*
	Jan Terje Bjørke	Properties of Networks to be Considered in their Visualisation *(Presentation) (Paper)*
1200	**Lunch**	
1300	Zack Jacobson	Introduction to breakout groups; topic selection
1330	Tour of Radar Site	
1430	**Provocations and Discussions**	
	Joanne Treurniet	*The Cyber Social Network Analogy*
	Amy Vanderbilt	Custom Ontologies for Expanded Social Network Analysis *(Presentation) (Paper)*
	Sonya McMullen	Crossing the Longest Yard: Eight Strategies for Creating Knowledge from a Glut of Data *(Presentation) (Paper)*
1520	**Coffee Break**	
1530	**Breakout Groups – First Working Session**	
1715	End, Day 1	

Agenda

Wednesday, 19th October 2005

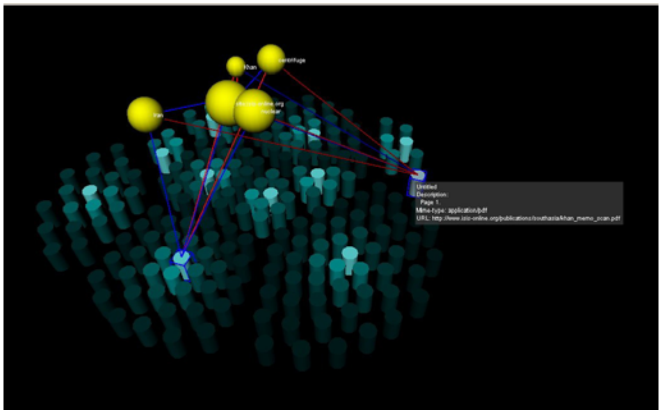
0930	Zack Jacobson	Good Morning
0945	**Provocations and Discussions**	
	Ben Houston	*Exploratory Visualization of Infectious Disease Propagation*
	Annette Kaster	*Network Visualisation of Object Relations Extracted from J2 Messages*
	Margaret Varga	*Infectious Disease Intervention Management*
1045	**Breakout Groups – Second Working Session**	
1215	**Lunch**	
1315	**Plenary Session**	
	Breakout Presentations	
	Mark Nixon, Martin Taylor, Margaret Varga, Jan Terje Bjørke, Amy Vanderbilt	Network Visualization: Reference Model *(Presentation) (Paper)*
	Lisbeth Rasmussen, Sonya McMullen	*Representing Uncertainty, Unknowns and Dynamics of Social Networks*
	Ben Houston, Marcus Lem, Vincent Taylor, Annette Kaster, Joanne Treurniet	*Unmasking a Terrorist*
1415	**Discussion**	
1500	**Coffee Break**	
1530	David Zeltzer, Marcus Lem	*Appreciation and Closing Discussion*
1630	Lisbeth Rasmussen	*Introducing 2006 Workshop, Invitation to Copenhagen*
1645	Zack Jacobson, Margaret Varga	Wrap-up
1700	Annette Kaster, FGAN	Adjournment, Farewell.

Annex L – 2007 VISUALISATION NETWORK-OF-EXPERTS WORKSHOP

Visualisation Network-of-Experts

Supporting NATO Research Task Group IST-059 on Network Visualisation

http://www.visn-x.net



Workshop:
Network Analysis and Visualisation for Simulation and Prediction

6th – 8th November 2007

The Aerospace Corporation
2350 E. El Segundo Blvd., El Segundo, CA 90245-4691, USA

Edited by M. Varga

The NATO Research Task Group “Visualisation Technologies for Network Analysis” has invited the Visualisation Network of Experts (Vis N/X) to advise it on how graph/network analysis and visualisation can support simulations of and predictions about dynamic systems critical to defence, intelligence and public safety.

L.1 WORKSHOP COMMITTEE

Dr. Mark Nixon (Aerospace Corp.)

Dr. Zack Jacobson (Health Canada)

Dr. Amy Vanderbilt (Vanderbilt Consulting)

Dr. Margaret Varga (QinetiQ)

L.2 WORKSHOP PARTICIPANTS

Visualisation Network of Experts Workshop; November 6-8, 2007

Name	Affiliation/Organization	e-mail
Mark Nixon	Aerospace Corp	mark.r.nixon@aero.org
Margaret Varga	QINETIQ	mjvarga@mail.qinetq.com
Amy KCS Vanderbilt	Vanderbilt Consulting	avanderbilt@vanderbilt-consulting.com
Andy Swarbrick	DETICA	andy.swarbrick@detica.com
Joanne Treurniet	DRDC	joanne.treurniet@drdc-rddc.gc.ca
Marcus Lem	Public Health Agency of Canada	marcus_lem@phac-aspc.gc.ca
Rob Young	DSTL	riyoung@dstl.gov.uk
Zack Jacobson	Health Canada	zack_jacobson@hc-sc.gc.ca zack@sigmaxi.net
Martin Taylor	Martin Taylor Consulting	mmt@mmtaylor.net
Frank Meng	Aerospace Corp	frank.meng@aero.org
Alain Bouchard	DRDC Valcartier	alain.bouchard@drdc-rddc.gc.ca
Donna Nystrom	Aerospace Corp	donna.m.nystrom@aero.org
Jan Terje Bjørke	FFI Norway	jtb@ffi.no
Lisbeth M. Rasmussen	DALO	lr@mil.dk
Vincent Taylor	DRDC Ottawa	vincent.taylor@drdc-rddc.gc.ca
David Hall	Penn State University	dhall@ist.psu.edu
Cristin Hall	Tech Reach Inc.	cmh187@psu.edu
Edward Palazzolo	Arizona State University	etp@asu.edu

CTRL, Click on the *blue, italicised* links to view the associated files.

Agenda

Tuesday, 6th November 2007

0845	Security check-in	
0915	Mark Nixon	*Welcome to Aerospace*
0920		Information on local arrangements
0930	Round Robin Introductions	
0945	Vincent Taylor	*RTG025 and the N/X*
1000	Zack Jacobson, Margaret Varga	Workshop Introduction
1015	**Coffee Break**	
1030	**Provocations and Discussions**	
	Amy K.C.S. Vanderbilt	Predictability of Dynamic Network Behavior – Unanswered Questions and Possible Directions *(Abstract) (Presentation)*
	Zack Jacobson	*VITA – Visual Interface for Text Analysis*
	Dragos Calitoiu	Knowledge Discovery in a Two-Mode Network: A Visualization Approach to Measure Interdependence between Actors and Social Venues *(Paper) (Presentation)*
	Dragos Callitoiu, Zachary Jacobson	Measuring Information Integration in Social Networks *(Paper provided but not presented)*
1145	**Lunch**	
1245	**Provocations and Discussions**	
	Donna Nystrom	Social Network Analysis & Link Discovery of Florida Department of Corrections Data *(Abstract) (Presentation)*
	Margaret Varga, Jan Terje Bjørke	*Hyper-Network with Uncertainties*
	Jan Terje Bjørke, Margaret Varga	Algorithm to Construct Fuzzy Hyper-Networks *(Abstract) (Presentation)*
	Martin Taylor	*The VisTG Framework for Network Visualisation*
1345	**Introduction to Breakout Groups**	
	Zack Jacobson	**Topic Suggestions:** Painting Pictures of My Enemy: How Can I Model Networks without Empirical Data? Visualising Network Uncertainties – e.g., Structure, Traffic, Nodal Processing

Agenda

Zack Jacobson	**Topic Suggestions (cont'd):** Different Visualisations for Different Network Types? Visualization at the Edges of Chaos: How to Treat Approximate Network Prediction? … Uncertainty at Critical Junctions? Different Variables for Different Simulations: Can you Choose Variables to Visualise? How? Different Domains: Same Visualisation for Viruses and e-viruses? Testbed Datasets – How Do We Create Them? Wild Card: Choose Your Own Question?

1415 **Breakout Groups – First Working Session**

1715 End, Day 1

Agenda

Wednesday, 7th November 2007

0830	Security check-in	
0900	Amy K.C.S. Vanderbilt	Good Morning
0905	**Coffee Break**	
0930	**Provocations and Discussions**	
	Rob Young, Margaret Varga	Knowledge representation *(Abstract) (Presentation)*
	Margaret Varga, Rob Young	Visualisation in Intelligence *(Abstract) (Presentation)*
	Andy Swarbrick	Best Use of Quantitative Social Network Analysis, Measurement and Visualization *(Abstract)*
1045	**Breakout Groups – Second Working Session**	
1145	**Lunch**	
1245	**Provocations and Discussions**	
	Michael C. Otterstatter	How might we best test network models of disease spread? *(Abstract) (Presentation)*
	Dave Hall	*Use of Visualization Techniques to Support Cyber Situational Awareness*
	Ed Palazzolo	*Simulating Communication and Knowledge Networks as Transactive Memory Systems*
	Sonya A.H. McMullen, Margaret Varga, Cristin M. Hall	Integration of Advanced Forensics and Medical Information with Network Visualization Techniques for Rapid Disaster Response *(Abstract) (Presentation)*
1400	**Breakout Groups – Third Working Session. Presentation Preparations**	
1715	End, Day 2	
1900	No-host dinner	

Agenda

Thursday, 8th November 2007

0830	Security check-in	
0900	Zack Jacobson	Good morning
0905	**Breakout Presentations**	
	Martin Taylor, Mark Nixon, David Hall, Vincent Taylor	*Framework Development*
	Amy Vanderbilt, Marcus Lem, Cristin Hall, Joanne Treurniet, Rob Young	*Developing Network Testbed Data Sets*
	Alain Bouchard, Andy Swarbrick, Ed Palazzolo, Zack Jacobson, Donna Nystrom, Margaret Varga	*Visualisation of Uncertainties*
1030	**Workshop Conclusions and Recommendations**	
1100	Adjournment	
1130	Depart for JPL (optional)	
1300	JPL Tour	
1500	Tour ends	

Annex M – 2008 VISUALISATION NETWORK-OF-EXPERTS WORKSHOP

Visualisation Network-of-Experts

Supporting NATO Research Task Group IST-059 on Network Visualisation

http://www.visn-x.net

Priory Church, Malvern. 2008 © Martin Taylor

Workshop:
Visualising Network Dynamics

4th – 6th November 2008

QinetiQ Malvern
Malvern Technology Centre
St. Andrew's Road, Malvern, Worcestershire, WR14 3PS, GBR

Edited by M. Varga

The NATO Research Task Group "Visualisation Technologies for Network Analysis" has invited the Visualisation Network of Experts (Vis N/X) to advise it on how graph/network analysis and visualisation can support simulations of and predictions about dynamic systems critical to defence, intelligence and public safety.

CTRL, Click on the *blue*, *italicised* links to view the associated files.

Agenda

Tuesday, 4th November 2008

0945	Security check-in	
1015	Margaret Varga	Welcome to QinetiQ
1030		Information on local arrangements
1035	Round Robin Introductions	
1055	Vincent Taylor	Introduction to NATO IST-059 and the N/X
1105	Zack Jacobson, Margaret Varga	Workshop Introduction
1115	**Provocations and Discussions: Design and Evaluation**	
	Amy K.C.S. Vanderbilt	*Evaluation of Interactive Visualisation of Network Dynamics*
	Rob Young	Some Considerations towards Metrics to Assess Visualisation *(Abstract) (Presentation)*
	Marcus Lem, Ben Houston	*Design and Specifications of a Visualization Tool for Dynamic Network Analysis (VITA-DNA)*
1215	**Lunch**	
1300	**Provocations and Discussions: Techniques and Applications**	
	Neil Bowan	CBRN Incident and Consequence Management and CBRN Capability Policy
	Nisha Iswaran	Effects-Based Approach to CBRN Capability Policy
1340	**Provocations and Discussions: Techniques and Applications**	
	Sven Brueckner	*Entities: An Alternative Perspective on Multiple Networks*
	Rusty Bobrow	*Kinetic Visualizations: Seeing and Understanding Structure in Large Interacting Networks, Using Human Motion Perception*
	Dragos Calitoiu, Zack Jacobson, Margaret Varga, Ben Houston	*Knowledge Discovery in a Multi-Mode Network: A Visualization Approach to Measure the Interdependence between Actors and Social Venues*
	Jan Terje Bjørke, Margaret Varga	*Weight (Error) Propagation in Networks of Hypernodes*
1500	**Introduction to Breakout Groups**	
	Zack Jacobson	Topic Selection
1510	**Breakout Groups – First Working Session**	
1630	End, Day 1	
1900	Dinner at the Red Lion Pub	

Agenda

Wednesday, 5th November 2008

0820	Security check-in	
0830	Margaret Varga	Welcome
0845	**Provocations and Discussions: Techniques and Applications (I)**	
	Martin Taylor	The IST-059 Framework for Network Visualisation *(Presentation) (Paper)*
	Sonya McMullen, Cristin Hall	Survey of Relevant Research Related to Data Visualization Applications and Areas Requiring Additional Study *(Presentation) (Paper)*
	Andy Swarbrick	*"Can you evaluate a tool to visualise a complex system, like an organisational network, without validated models for how it behaves?"– Drawing the lines between exploration and discovery versus understanding and prediction. Why we need to work anthropologists and social scientists.*
0940	**Breakout Groups – Second Working Session**	
1115	**Provocation and Discussions: Techniques and Applications (II)**	
	Shashi Shekhar	*Spatio-Temporal Networks: A GIS Perspectives*
	Paul Wonnacott	*Geo-Temporal Visualisation of Networks for Intelligence*
	Neil Briscombe	Application of Semantic Tooling for Military Information Exploitation and Decisions Support Systems
1215	**Lunch**	
1245	**Provocations and Discussions: Techniques and Applications (III)**	
	Zack Jacobson, Ben Houston	*VITA – Visual Interface for Text Analysis*
	Michael Otterstatter, Zack Jacobson	*Contagion in Real Social Networks: Insights from Social Insects*
1340	**Breakout Groups – Third Working Session**	
1545	End, Day 2	
1600	Depart for wine tasting event	
1900	Dinner at the ASK restaurant	

Agenda

Thursday, 6th November 2008

0845	Security check-in	
0900	Amy K.C.S. Vanderbilt	Welcome
0910	**Provocations and Discussions: Techniques and Applications (IV)**	
	Andy Swarbrick	SNA Visualisation Tools
	Dr. Mark Round	A Research Agenda for SNA
	Margaret Varga, Dr. Rob Young, Kevin Adams and Rose Hines	*Hypothesis Network Visualisation*
1010	**Breakout Groups: Final Working Session and Preparation of Presentations**	
1215	**Lunch**	
1300	**Breakout Group Presentations: Workshop Conclusions and Recommendations**	
	Cristin Hall, Chris Horeczy, Annette Kaster, Michael Kleiber, Marcus Lem, Sonya McMullen	*Experimental Designs for Utility Testing of Visualization Techniques*
	Neil Bowman, James Freemantle, Rose Hines, Nisha Iswaran, Andy Swarbrick, Margaret Varga, Rob Young	*Visualising the Propagation of Uncertainties*
	Amy K.C.S. Vanderbilt, Vincent Taylor, Martin Taylor, Mark Nixon, Jan Terje Bjørke, Sven Brueckner, Zack Jacobson, Jason Moore, Rusty Bobrow	*Information Theoretic Considerations*
	Kevin Adams, Mark Round, Andrew Webb, Lisbeth Ramussen	*Approaches to the Analysis and Visualization of Multi-Modal and Multi-Relational Networks*
1450	Margaret Varga	Close of Workshop and Farewell
1500	Workshop Closes	

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's References	3. Further Reference	4. Security Classification of Document
	RTO-TR-IST-059 AC/323(IST-059)TP/262	ISBN 978-92-837-0084-5	UNCLASSIFIED/ UNLIMITED

5. Originator

Research and Technology Organisation
North Atlantic Treaty Organisation
BP 25, F-92201 Neuilly-sur-Seine Cedex, France

6. Title

A Framework for Network Visualisation

7. Presented at/Sponsored by

Final Report of Task Group IST-059.

8. Author(s)/Editor(s)	9. Date
Multiple	February 2010

10. Author's/Editor's Address	11. Pages
Multiple	374

12. Distribution Statement

There are no restrictions on the distribution of this document. Information about the availability of this and other RTO unclassified publications is given on the back cover.

13. Keywords/Descriptors

Command and control
Computer networks
Counterterrorism
Data overload
Fuzzy networks
Human-machine interface
Hypernode
Information theory
Information visualisation
Knowledge discovery
Link analysis
Massive datasets
Multimedia visualisation
Network analysis
Network centric warfare
Network dynamics
Network theory
Situation assessment
Social networks

14. Abstract

The programme and findings of the NATO RTO Technical Team IST-059 "A Framework for Network Visualisation" are described in a report divided into nine chapters, supplemented by technical annexes. The team developed a comprehensive framework for network visualisation (Chapter 2); undertook a survey of network visualisation tools and methods (Chapter 3); developed an information-theoretic tool (the hypernode) for network analysis (Chapter 4); defined an interface between the survey and the framework, to aid the identification of visualisation technologies that match a particular domain space and user role (Chapter 5); walked through test cases from different domains to validate the concept and design of the framework and to provide insight into the linkages of Chapter 5 (Chapter 6); sponsored annual workshops in support of the programme (Chapter 7), addressing the following topics of special interest: Social Network Analysis and Visualisation for Public Safety; Visualising Network Information; Network Analysis and Visualisation for Simulation and Prediction; and Visualising Network Dynamics. An unexpected and surprising outcome of the programme's information-theoretic research and conceptual framework development was the observation that together they appear to provide a starting point from which a "Unified Theory of Networks" may be developed.

NORTH ATLANTIC TREATY ORGANISATION

RESEARCH AND TECHNOLOGY ORGANISATION

BP 25
F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE
Télécopie 0(1)55.61.22.99 • E-mail mailbox@rta.nato.int

DIFFUSION DES PUBLICATIONS
RTO NON CLASSIFIEES

Les publications de l'AGARD et de la RTO peuvent parfois être obtenues auprès des centres nationaux de distribution indiqués ci-dessous. Si vous souhaitez recevoir toutes les publications de la RTO, ou simplement celles qui concernent certains Panels, vous pouvez demander d'être inclus soit à titre personnel, soit au nom de votre organisation, sur la liste d'envoi.

Les publications de la RTO et de l'AGARD sont également en vente auprès des agences de vente indiquées ci-dessous.

Les demandes de documents RTO ou AGARD doivent comporter la dénomination « RTO » ou « AGARD » selon le cas, suivi du numéro de série. Des informations analogues, telles que le titre est la date de publication sont souhaitables.

Si vous souhaitez recevoir une notification électronique de la disponibilité des rapports de la RTO au fur et à mesure de leur publication, vous pouvez consulter notre site Web (www.rto.nato.int) et vous abonner à ce service.

CENTRES DE DIFFUSION NATIONAUX

ALLEMAGNE
Streitkräfteamt / Abteilung III
Fachinformationszentrum der Bundeswehr (FIZBw)
Gorch-Fock-Straße 7, D-53229 Bonn

BELGIQUE
Royal High Institute for Defence – KHID/IRSD/RHID
Management of Scientific & Technological Research
for Defence, National RTO Coordinator
Royal Military Academy – Campus Renaissance
Renaissancelaan 30, 1000 Bruxelles

CANADA
DSIGRD2 – Bibliothécaire des ressources du savoir
R et D pour la défense Canada
Ministère de la Défense nationale
305, rue Rideau, 9e étage
Ottawa, Ontario K1A 0K2

DANEMARK
Danish Acquisition and Logistics Organization (DALO)
Lautrupbjerg 1-5, 2750 Ballerup

ESPAGNE
SDG TECEN / DGAM
C/ Arturo Soria 289
Madrid 28033

ETATS-UNIS
NASA Center for AeroSpace Information (CASI)
7115 Standard Drive
Hanover, MD 21076-1320

FRANCE
O.N.E.R.A. (ISP)
29, Avenue de la Division Leclerc
BP 72, 92322 Châtillon Cedex

GRECE (Correspondant)
Defence Industry & Research General
Directorate, Research Directorate
Fakinos Base Camp, S.T.G. 1020
Holargos, Athens

HONGRIE
Department for Scientific Analysis
Institute of Military Technology
Ministry of Defence
P O Box 26
H-1525 Budapest

ITALIE
General Secretariat of Defence and
National Armaments Directorate
5th Department – Technological
Research
Via XX Settembre 123
00187 Roma

LUXEMBOURG
Voir Belgique

NORVEGE
Norwegian Defence Research
Establishment
Attn: Biblioteket
P.O. Box 25
NO-2007 Kjeller

PAYS-BAS
Royal Netherlands Military
Academy Library
P.O. Box 90.002
4800 PA Breda

POLOGNE
Centralny Ośrodek Naukowej
Informacji Wojskowej
Al. Jerozolimskie 97
00-909 Warszawa

PORTUGAL
Estado Maior da Força Aérea
SDFA – Centro de Documentação
Alfragide
P-2720 Amadora

REPUBLIQUE TCHEQUE
LOM PRAHA s. p.
o. z. VTÚLaPVO
Mladoboleslavská 944
PO Box 18
197 21 Praha 9

ROUMANIE
Romanian National Distribution
Centre
Armaments Department
9-11, Drumul Taberei Street
Sector 6
061353, Bucharest

ROYAUME-UNI
Dstl Knowledge and Information
Services
Building 247
Porton Down
Salisbury SP4 0JQ

SLOVAQUIE
Akadémia ozbrojených síl
M.R. Štefánika, Distribučné a
informačné stredisko RTO
Demanova 393, P.O.Box 45
031 19 Liptovský Mikuláš

SLOVENIE
Ministry of Defence
Central Registry for EU and
NATO
Vojkova 55
1000 Ljubljana

TURQUIE
Milli Savunma Bakanlığı (MSB)
ARGE ve Teknoloji Dairesi
Başkanlığı
06650 Bakanliklar
Ankara

AGENCES DE VENTE

**NASA Center for AeroSpace
Information (CASI)**
7115 Standard Drive
Hanover, MD 21076-1320
ETATS-UNIS

**The British Library Document
Supply Centre**
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
ROYAUME-UNI

**Canada Institute for Scientific and
Technical Information (CISTI)**
National Research Council Acquisitions
Montreal Road, Building M-55
Ottawa K1A 0S2, CANADA

Les demandes de documents RTO ou AGARD doivent comporter la dénomination « RTO » ou « AGARD » selon le cas, suivie du numéro de série (par exemple AGARD-AG-315). Des informations analogues, telles que le titre et la date de publication sont souhaitables. Des références bibliographiques complètes ainsi que des résumés des publications RTO et AGARD figurent dans les journaux suivants :

Scientific and Technical Aerospace Reports (STAR)
STAR peut être consulté en ligne au localisateur de ressources
uniformes (URL) suivant: http://www.sti.nasa.gov/Pubs/star/Star.html
STAR est édité par CASI dans le cadre du programme
NASA d'information scientifique et technique (STI)
STI Program Office, MS 157A
NASA Langley Research Center
Hampton, Virginia 23681-0001
ETATS-UNIS

Government Reports Announcements & Index (GRA&I)
publié par le National Technical Information Service
Springfield
Virginia 2216
ETATS-UNIS
(accessible également en mode interactif dans la base de
données bibliographiques en ligne du NTIS, et sur CD-ROM)

NORTH ATLANTIC TREATY ORGANISATION

RESEARCH AND TECHNOLOGY ORGANISATION

BP 25
F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE
Télécopie 0(1)55.61.22.99 • E-mail mailbox@rta.nato.int

DISTRIBUTION OF UNCLASSIFIED RTO PUBLICATIONS

AGARD & RTO publications are sometimes available from the National Distribution Centres listed below. If you wish to receive all RTO reports, or just those relating to one or more specific RTO Panels, they may be willing to include you (or your Organisation) in their distribution.

RTO and AGARD reports may also be purchased from the Sales Agencies listed below.

Requests for RTO or AGARD documents should include the word 'RTO' or 'AGARD', as appropriate, followed by the serial number. Collateral information such as title and publication date is desirable.

If you wish to receive electronic notification of RTO reports as they are published, please visit our website (www.rto.nato.int) from where you can register for this service.

NATIONAL DISTRIBUTION CENTRES

BELGIUM
Royal High Institute for Defence – KHID/IRSD/RHID
Management of Scientific & Technological Research
for Defence, National RTO Coordinator
Royal Military Academy – Campus Renaissance
Renaissancelaan 30
1000 Brussels

CANADA
DRDKIM2 – Knowledge Resources Librarian
Defence R&D Canada
Department of National Defence
305 Rideau Street, 9th Floor
Ottawa, Ontario K1A 0K2

CZECH REPUBLIC
LOM PRAHA s. p.
o. z. VTÚLaPVO
Mladoboleslavská 944
PO Box 18
197 21 Praha 9

DENMARK
Danish Acquisition and Logistics Organization (DALO)
Lautrupbjerg 1-5
2750 Ballerup

FRANCE
O.N.E.R.A. (ISP)
29, Avenue de la Division Leclerc
BP 72, 92322 Châtillon Cedex

GERMANY
Streitkräfteamt / Abteilung III
Fachinformationszentrum der Bundeswehr (FIZBw)
Gorch-Fock-Straße 7
D-53229 Bonn

GREECE (Point of Contact)
Defence Industry & Research General Directorate
Research Directorate, Fakinos Base Camp
S.T.G. 1020
Holargos, Athens

HUNGARY
Department for Scientific Analysis
Institute of Military Technology
Ministry of Defence
P O Box 26
H-1525 Budapest

ITALY
General Secretariat of Defence and
National Armaments Directorate
5th Department – Technological
Research
Via XX Settembre 123
00187 Roma

LUXEMBOURG
See Belgium

NETHERLANDS
Royal Netherlands Military
Academy Library
P.O. Box 90.002
4800 PA Breda

NORWAY
Norwegian Defence Research
Establishment
Attn: Biblioteket
P.O. Box 25
NO-2007 Kjeller

POLAND
Centralny Ośrodek Naukowej
Informacji Wojskowej
Al. Jerozolimskie 97
00-909 Warszawa

PORTUGAL
Estado Maior da Força Aérea
SDFA – Centro de Documentação
Alfragide
P-2720 Amadora

ROMANIA
Romanian National Distribution
Centre
Armaments Department
9-11, Drumul Taberei Street
Sector 6, 061353, Bucharest

SLOVAKIA
Akadémia ozbrojených síl
M.R. Štefánika, Distribučné a
informačné stredisko RTO
Demanova 393, P.O.Box 45
031 19 Liptovský Mikuláš

SLOVENIA
Ministry of Defence
Central Registry for EU & NATO
Vojkova 55
1000 Ljubljana

SPAIN
SDG TECEN / DGAM
C/ Arturo Soria 289
Madrid 28033

TURKEY
Milli Savunma Bakanlığı (MSB)
ARGE ve Teknoloji Dairesi
Başkanlığı
06650 Bakanliklar – Ankara

UNITED KINGDOM
Dstl Knowledge and Information
Services
Building 247
Porton Down
Salisbury SP4 0JQ

UNITED STATES
NASA Center for AeroSpace
Information (CASI)
7115 Standard Drive
Hanover, MD 21076-1320

SALES AGENCIES

NASA Center for AeroSpace Information (CASI)
7115 Standard Drive
Hanover, MD 21076-1320
UNITED STATES

The British Library Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
UNITED KINGDOM

Canada Institute for Scientific and Technical Information (CISTI)
National Research Council Acquisitions
Montreal Road, Building M-55
Ottawa K1A 0S2, CANADA

Requests for RTO or AGARD documents should include the word 'RTO' or 'AGARD', as appropriate, followed by the serial number (for example AGARD-AG-315). Collateral information such as title and publication date is desirable. Full bibliographical references and abstracts of RTO and AGARD publications are given in the following journals:

Scientific and Technical Aerospace Reports (STAR)
STAR is available on-line at the following uniform resource
locator: http://www.sti.nasa.gov/Pubs/star/Star.html
STAR is published by CASI for the NASA Scientific
and Technical Information (STI) Program
STI Program Office, MS 157A
NASA Langley Research Center
Hampton, Virginia 23681-0001
UNITED STATES

Government Reports Announcements & Index (GRA&I)
published by the National Technical Information Service
Springfield
Virginia 2216
UNITED STATES
(also available online in the NTIS Bibliographic Database
or on CD-ROM)

ISBN 978-92-837-0084-5