Modelling and Simulation Process in the Polish Proposal of Early Warning System for DAT and Crisis Management

Andrzej Najgebauer PhD, DSc, Prof. of MUT
Military University of Technology,
Faculty of Cybernetics
2 Kaliskiego Str.
00-908 Warsaw
Poland

e-mail: anajgeb@isi.wat.edu.pl

ABSTRACT

An approach to the Early Warning Systems building is proposed. The process of modelling the early identification processes is presented on the basis of MSG026 group experiences. The idea of the terrorist threat pattern acquisitionis developed. Some data mining techniques including filtering, semantic nets, link analysis, neural networks, discriminant analysis in the process of threat recognition were disscussed and modelled . The architecture of the Early Warning System was proposed in the three-layer J2EE approach - client layer, application layer and database system. One of the most interesting features of the proposed architecture there is possibility of dynamic functionality enhancement. Some experiments in the environmente were conducted on the basis of training data set in the part of EWS – expert Corvid system and interactive simulator of terrorist attack against an infrastructure as the demonstrator tools of potential complex analysis and prediction. The expert system was build on the basis of terrorist attack scenario set.

Najgebauer, A. (2005) Modelling and Simulation Process in the Polish Proposal of Early Warning System for DAT and Crisis Management. In *The Effectiveness of Modelling and Simulation – From Anecdotal to Substantive Evidence* (pp. 2-1 – 2-20). Meeting Proceedings RTO-MP-MSG-035, Paper 2. Neuilly-sur-Seine, France: RTO. Available from: http://www.rto.nato.int/abstracts.asp.

1.0 INTRODUCTION

Concerning to the study MSG 026 we would like to present a progress of modelling, simulation and development of an Early Warning System for diagnosis and signalling terrorist activities. Our publication entitled "A Concept of Simulation Based Diagnostic Support Tool for Terrorism Threat Awarenes" [1] were presented during the previous NMSG Conference in Koblenz. We have defined basic concepts, terms and models that enable complex approach to the proper identification and simulation terrorist activities and the consequences. It was proposed, that Early Warning System can be considered as a part of Crisis Management System and understood as information system that processes any information from any source about escalatory developments, be they slow and gradual or quick and sudden, far enough in advance in order for a national government or an international or regional organisation to react timely and effectively, if possible still leaving them time to employ preventive measures [7].

In a sense the early warning system (EWS) is a simulation-based diagnostic support tool with its associated algorithms that realises the following processes [6]:

- collecting information relevant to the terrorism threat estimation and intelligence data analysis from:
  - primary threat factors determination
  - aggregated threat factors (causative and executive) determination
  - threat coefficient estimation
  - possible goals of terrorist attack identification
- the analysis and simulation of the information in order to predict terrorism threat over long periods of time, the stability of the threat factors and the signalling when a break-through of a pre-determined threshold is detected.
- the visualisation of EWS output for potential users.

The ideas and models presented in the paper, there are results of the common work of MSG 026 group, especially Polish members of the group. However the most of the concepts were discussed within whole international group during our meetings.

2.0 DESCRIPTION OF THREAT ASSESSMENT

In the paper we would like to discuss some advances of MSG026 group in the problem of constructing the information framework. The framework should be heterogeneous environment for interactive simulation, which perform the processes of EWS. The scheme of the analysis can be presented like this:

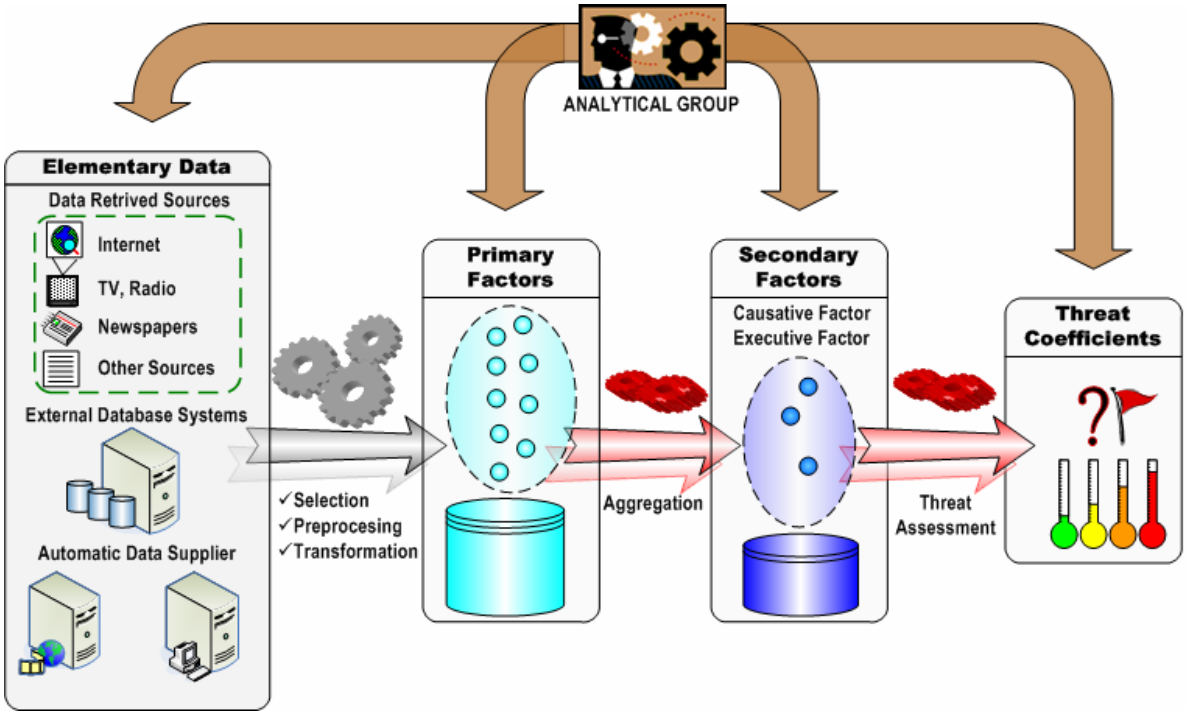


Fig. 1. Threat assessment process [1].

In the process of threat assessment we have used the aggregation of data, which is normally applied by analyst in the evaluation of real threat.

The flexible solution is required that process combines 3 phases [1]:

1.The transformation the elementary data into primary data factors.

2.The aggregation of primary factors into secondary factors (small set).

3.The transformation of secondary factors into threat coefficient.

Complete automation can include phases of elements data transformation – there is rather calculation of primary factors value on the basis of elementary data verified and validated. Next phases there are pattern recognition on the basis of historical data introduced by experts (case studies of difficult crisis) and threat assessment on the basis of recognized set of factors.

In the process of information model construction we have proposed two sets of Primary and Secondary Factors (fig. 2).

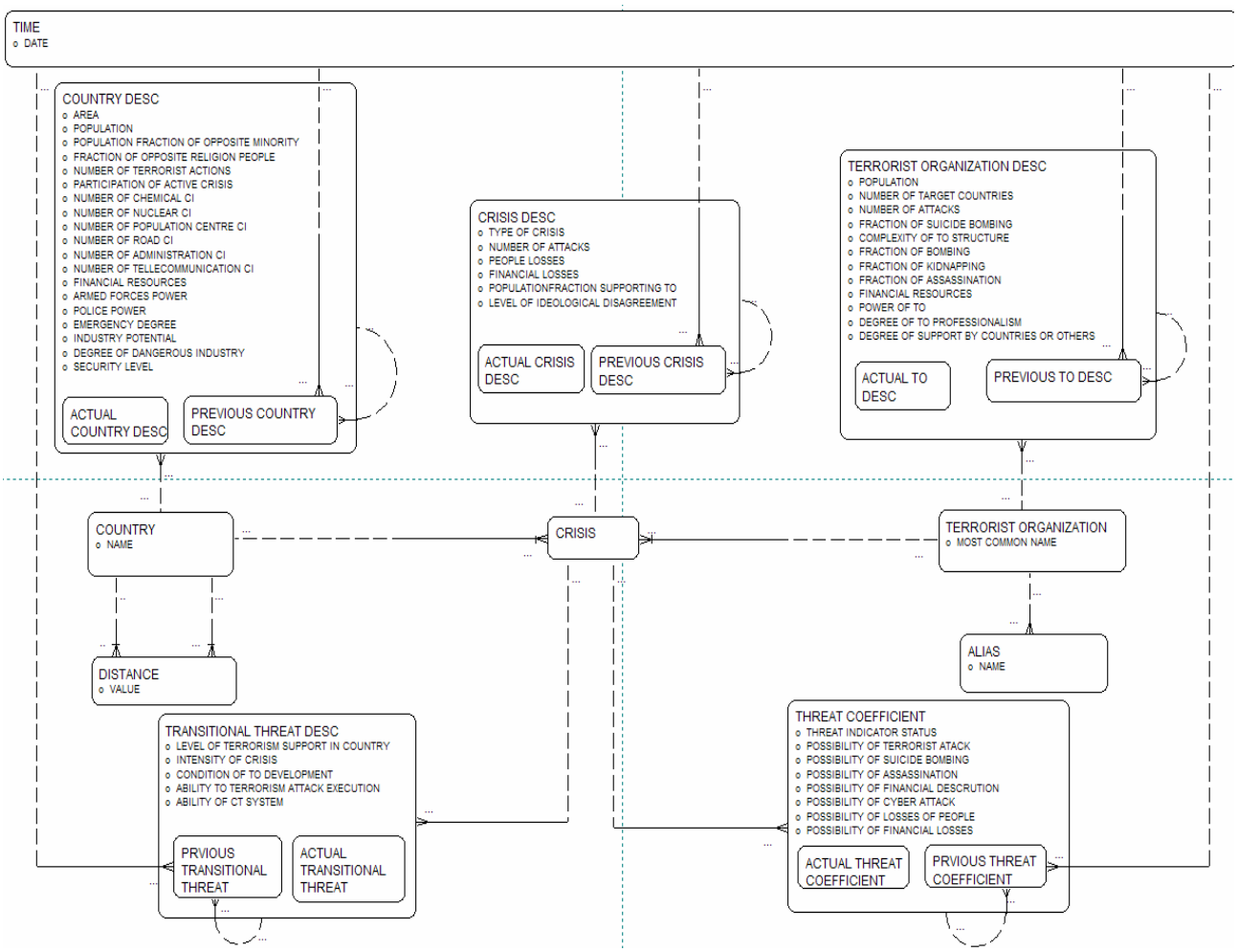


Fig.2. The information model of primary and secondary factors [4].

There are several concepts of acquisition the terrorist activities data for the intelligence, where further they can be processed by data mining mechanisms. Scanning large amount of data is the main task in many EWS (Early Warning Systems). The members of MSG026 group proposed some approaches to the scanning process [2], [3], [4]. This data could be represented in many different forms, one of them are semantic networks. This tool gives one of the most important advantages for such representation – scalibility and flexibility of knowledge representation. Presented method of semantic network analysis and association acquiring, aims at:

- Providing a tool for operating on large information resources,
- Eliminating the unreliable and unwanted information within the semantic network (essential requirement due to algorithm complexity),
- Selecting significant nodes and relations between them (for the analysis),
- Searching the indirect relations between the nodes based on already stored knowledge in the semantic network (building the new knowledge in the system).

The concept of building such knowledge base and most of all providing mechanisms for analyzing the data placed in it, consists of two phases:

1. Designing the set of ontologies used as filters for separating the most acurate and needed data at the moment of the analysis.
2. Providing a set of mesures for evaluating the correctness of the stored knowledge, and most of all, pointing the nodes which are suspected of indirect association.

The semantic network and ontologies will be implemented using the RDF and OWL languages along with the inference engine which will be constructed using the JENA framework.

The Semantic graph consists of nodes and directed connections between them. Each node represents a type of (e.g. person, event, object).

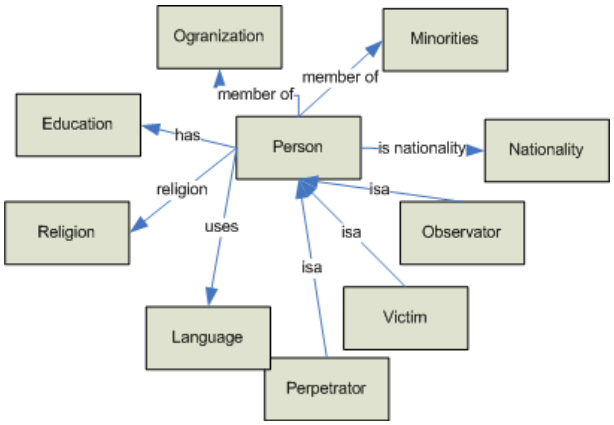


Fig. 3 An example of ontology for the semantic network

To represent specific associations between nodes there had been introduced several predefined assosciation types: is-a and part-of (representing the concepts of subtypes - inherritance and aggregation in object oriented approach). Ontology could be described as a template (metadata), which controlls the process of adding data to semantic network by allowing certain types of nodes to be connected by certain types of association [2].

Main problem while building semantic network is to balance the flexability of data representation and the information detail stored in the network. It is worth to mention that semantic networks are able to store the information of the type of node and association along with the detailed data. Such architecture forces to distinguish between certain types of information stored in the network due to the proper association aquisition (network path finding algorithm).

Semantic networks were developed as a method of representing large scale data sets. The analysis of such tool is very complex. There are several problems to overcome: one of them is the ambiguous meaning of the associations between different types of nodes, which may cause misleading interpretation of such association (incorrect results of path finding algorythm). The other one is the complexity of algorithm while considering several meanings of similar associations (similar names of stored relations between nodes). Presented method of association building in semantic networks uses two different methods of semantic network size reduction. The first one are additional ontologies used to filter the information in the semantic network. Ontology as a template (metamodel) is able to classify, nodes which belong to the same type and

which types of association can exitst between them. Application of such filter discards needless information, leaving only nodes and associations which will be taken into consideration by the association acquisition algorithms.

Other way to solve the coplexity problem is to introduce the quantitative analysis of the stored information. Introducing the set of coefficiets used to evaluate graphs or networks can enables to calculate several parameters describing the proper construction of the semantic network. Such analysis subjects to both nodes and associations, considering kwnoledge stored in semantic network as also as in ontology.

3.0. THE CONCEPT OF ASSOCIATION AQUISITION

Introduced method of building associations contains several stages [2]:

- Designing generic dynamicly changing ontology allowing flexible information representation;
- Designing the mechanisms for knowledge acquisition for building semantic network (also designing intelligent agents who provide the seeking database algorythms)
- Definition of parameters which will allow to analyze the knowlege evaluation in the network (dependency analysis, clustering connected to nodes concerning the base ontology)
- Definition of particular ontologies, which are used as filters for elimination of unnecessary links in the net.

To define specific paths lets introduce the following description[2]:

- „paths" - heuristic route search algorythms, are based on the problem size reduction using the reference to ontology analysis not the sematic network itself. Metamodel usage allows to decrese quantity of nodes and links to analyze by the algorythm. Linking nodes to calculated rote (building association) is being achieved using the depth-first algorythm, considering the currently analyzed node and the ontology templete wich gives the information of all valid links to other nodes;
- „intersecting paths" – using the definition of „ paths" on the semantic graph the algorithm is searching for two paths, containing intersecting links which connect nodes in those paths;
- „isomorfic paths" – are besed on the algorithm of finding two paths which are isomorfic, which means that taking two pairs of nodes we need to find such paths in the network that any of the nodes from one path is a part of the other path.

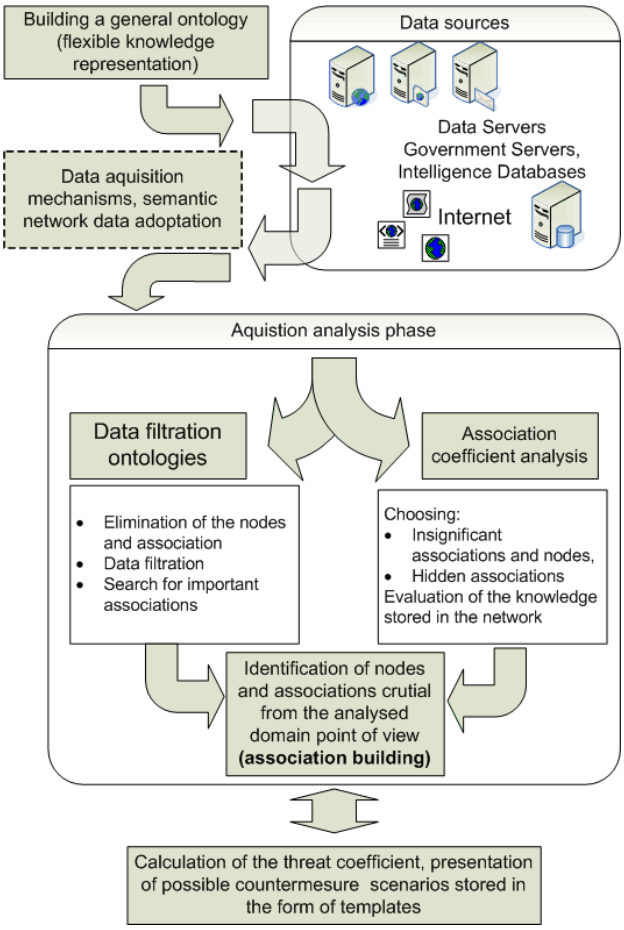


Fig. 4. Association building method for semantic networks [2]

4.0. AQUIRING NEW KNOWLEGDE

Aquiring new knowledge in semantic network is based on the introducing new nodes and links between them [2]. This can be achived in two ways: using the analysis od the structure of the semantic network, and inference engines. Inference algorithms, can be implemented as:

- general logic based inference engine – where there are two main aims, First Order Logic, Higher Order logic and Description Logic. First Order Logic (FOL) are machanisms which are very efficient but computationally not tractable for large amounts of data and axioms. Higher logic based engines however are able to track the inference route but they require a lot more resources to achive the task.
- solving algorithms – are specialized algorithms, often small size, designed to provide a solution in one distinct prolem. PSM (Problem Solving Methods) define which actions in the whole inference process need to be executed and how the controll flow in such algorithm should look like (considering the controll of the subtasks).

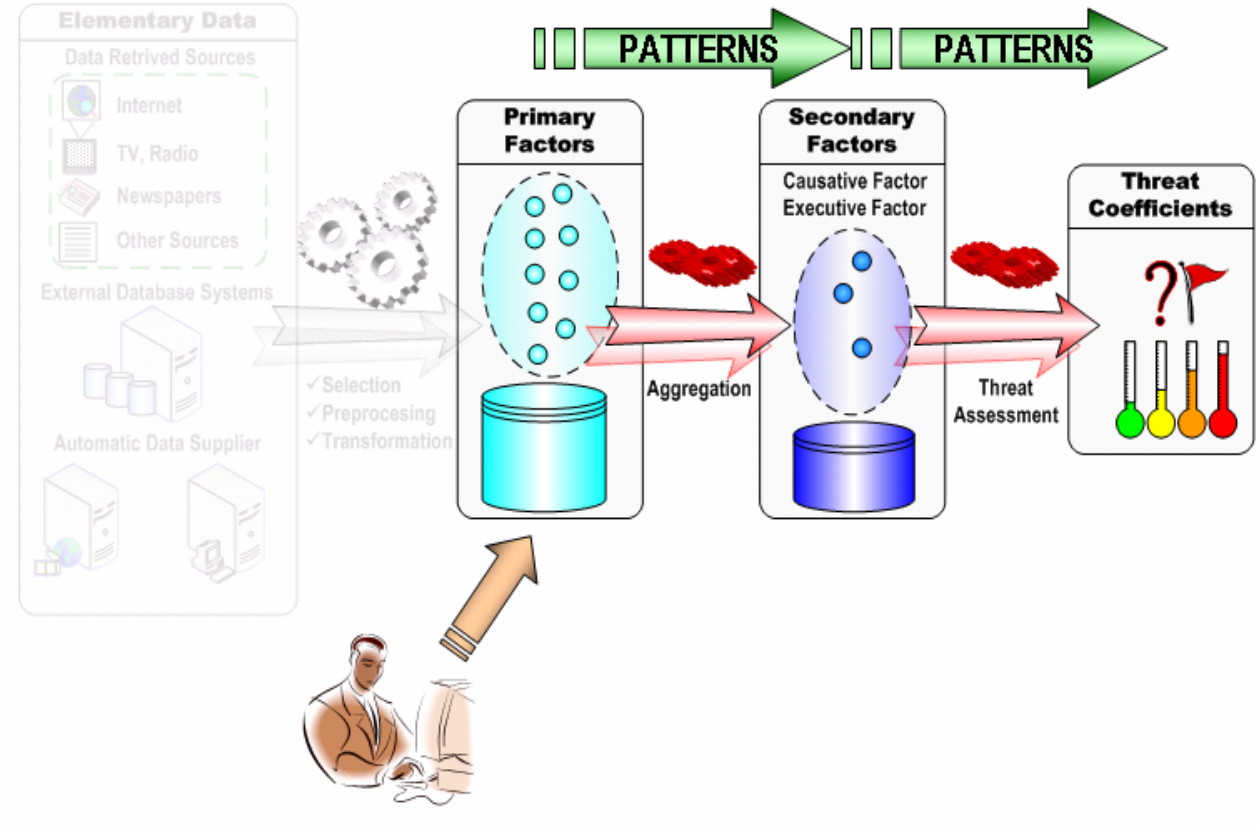


Fig.5. Training of the EWS – primary-secondary [1].

If we have prepared by experts training data set for primary factors in the description of entrants data and crisis, and values of secondary factors according to the situation described in primary factors for many period of time (real crisis – case studies) we can recognise a pattern of data transformation (aggregation).

For example primary factor: Armed_Forces_Power is relative value based on MOD budget, degree of modernity, degree of army training. All of these parameters should be calculated in relation to the greatest value probably USA. Such analysis should be provided for all primary factors considering wide spectrum of elementary data. These parameters, which aren't substantial for the threat assessment can be identify in the way of particular, statistical analysis.

As the secondary factor are concerned there is a classification proposed by intelligence community:

- Operational capability.
- Intension.

- Activity.
- Operational environment.

In the formal model these parameters are consider for the crisis, where we have minimum 2 sides: entrant threatened and terrorist organization.

Our proposal of secondary factors, which can be classified in the way above there are:

- Level of terrorist support in country.
- Intensity of crisis.
- Conditions of terrorist organization development.
- Ability to terrorist attack execution.
- Ability of CT system.

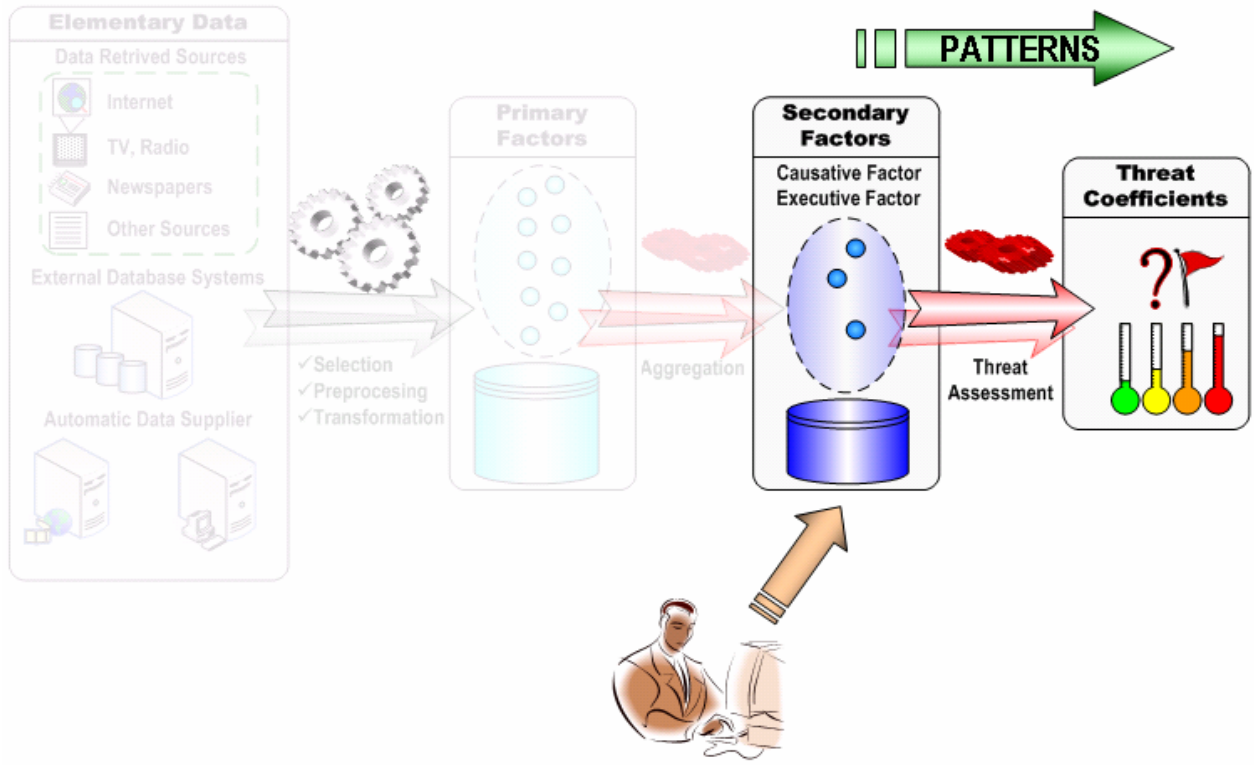
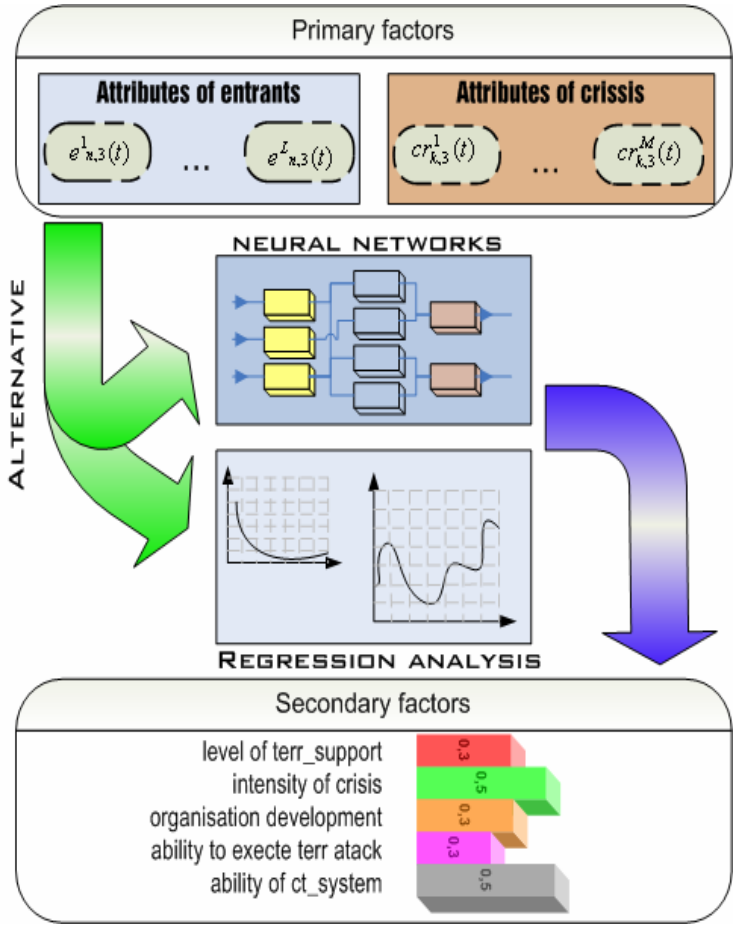


Fig.6. Training of the EWS – secondary-threat coefficient [1].

Analogous secondary factors input can be provided in the phase of determining of threat measure value and threat status. Having the pattern, which is recognised on the basis of earlier process of system learning.

The pattern recognition process can be provided by many known or completely new techniques:



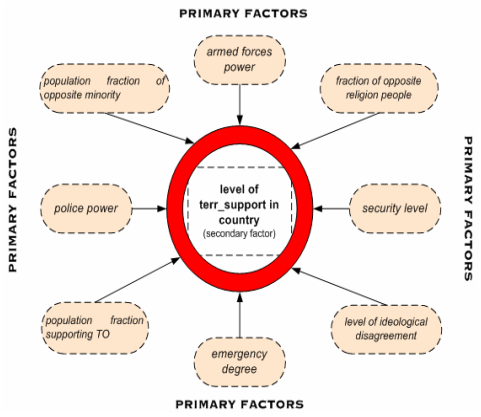


Fig. 7. Pattern recognition process

It is sometimes like backwards analysis. Doing it many times for different types of crisis (in a sense of causes and types of attack) we can learn our system (recognise the pattern) and use in the prediction process. We have proposed regression analysis and neural network multilayer perceptron. Another way, let say typical that is multivariate correlation analysis and step by step finding a functional dependence between secondary and primary factors. For example level_of_terr_support_in_country can depend on: population fraction of opposite minority, armed forces power, police power, security level, fraction of opposite religion people, population fraction supporting TO and so on. Adding new factors can enhance the set but doesn't break down the approach.

Using these two approaches we can calibrate our model and refine the recognised pattern in a figure of function or neural network structure and weight value (with hidden layers and determined set of neurons)(fig.8.). This type of neural network so called perceptron is good for discriminant analysis.

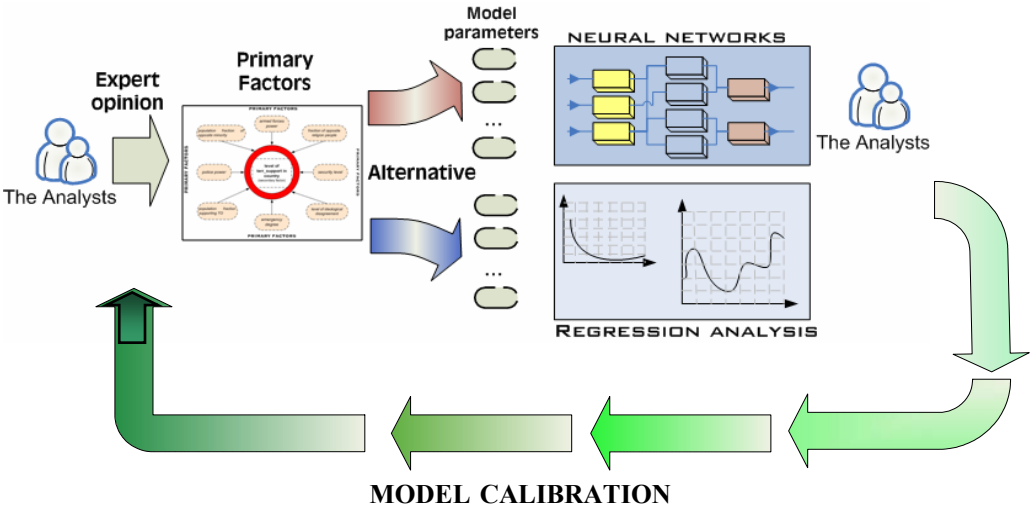


Fig.8. Calibration of the recognition tool

One of the verification methods of the presented approach we have proposed the clustering analysis for two levels of factors. In the analysis we can introduce many measures of classification for example as most typical there is Euclidean metrics in a sense geometric distance between object in a space or Chebychev metrics i.e. the shortest distance of maximum coordinate.

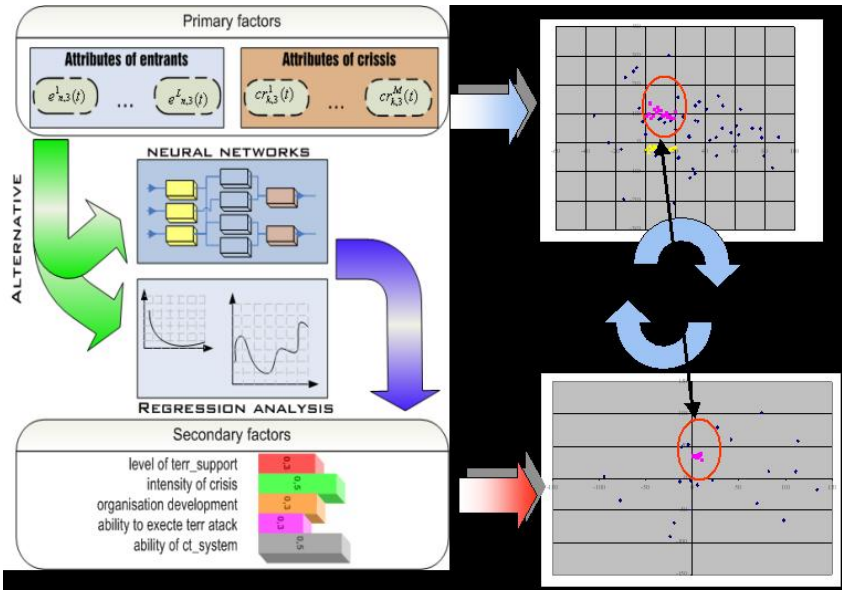


Fig.9. Clustering

5.0. ARCHITECTURE

The system architecture is based on MVC template and J2EE platform. It is known as distribute systems model architecture based on components and services, which are accessible in the platform.

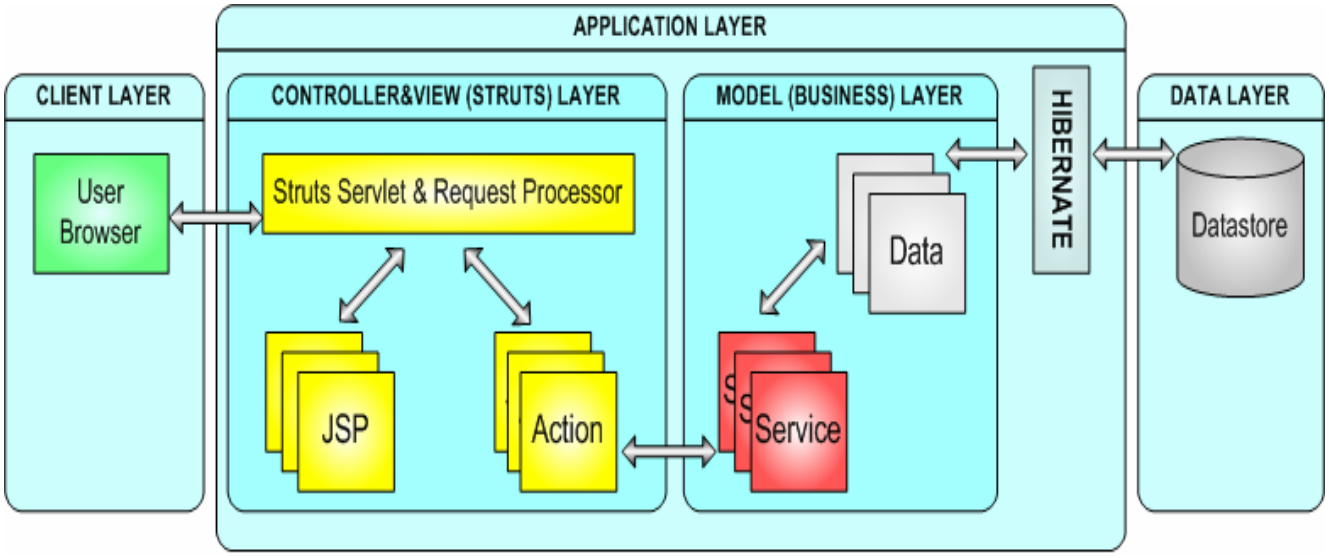


Fig.10. An Architecture of Early Warning System

The J2EE Application is usually three-layer:

1. Client layer responsible of system-user interaction. It is a thin client –HTML pages browser.
2. An application:
    a. Presentation layer responsible of pages generating for the thin client.
    b. Algorithms and control of computation (logic) – using EJB (Enterprise Java Beans) components.
3. Data base system.

One of the most important requirements in the simulation environment construction there is possibility of dynamic functionality enhancement without the general rebuilding of the environment. The requirement was fulfilled by using the components:

- public abstract String **describeMethod**();
- public abstract String **retrieveReturnType**();
- public abstract ArrayList<String> **retrieveParametersTypes**();

and then the introducing the enumerative types into Java 1.5.0 and the generic types.

These new functions are accessible for the client by using of RMI (*Remote Method Invocation*) technology.

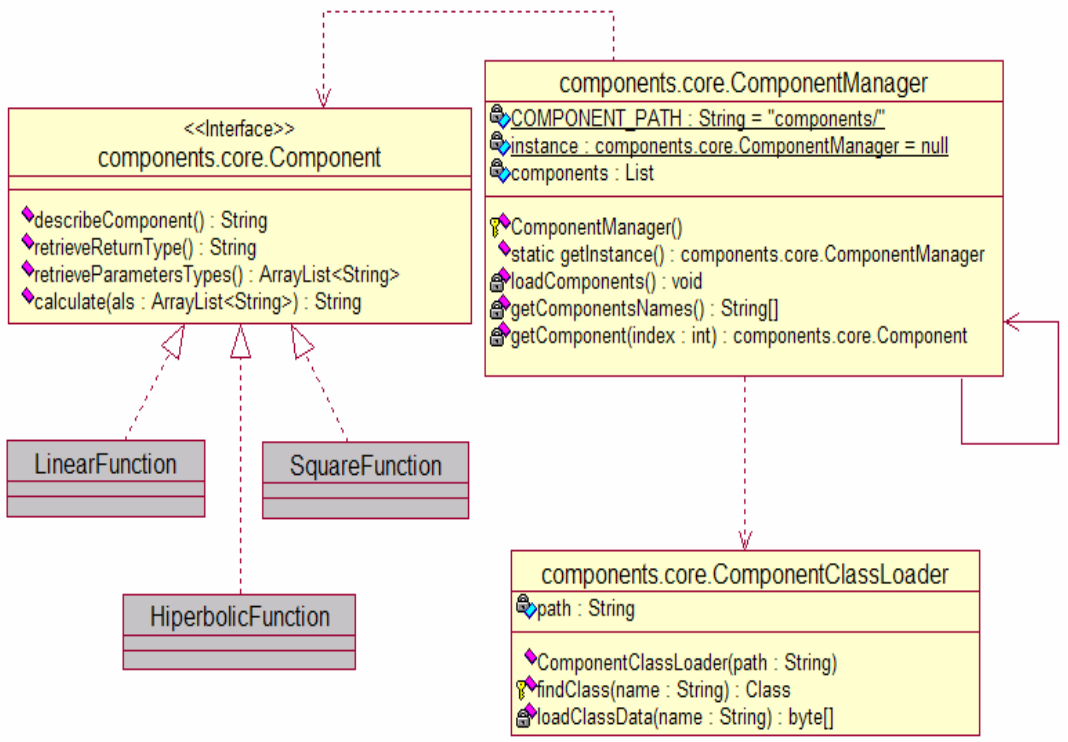


Fig. 11.An idea of dynamic functionality enhancement

The ClassLoader which is abstract class enables the construction of own ClassLoader in order to dynamic loading of new classes with previously unknown functionality.

6.0. IDENTIFICATION OF POSSIBLE METHODS OF TERRORIST ATTACKS – AN EXPERT SYSTEM SKELETON.

The skeletal expert system so called Exsys CORVID is produkt of Exsys Inc. It is very useful tool for rapid prototyping of expert system on the HTML pages. The results of resoning can be presented in the formatted way. Applet CORVID enables the functionality for most systems. As the data collecting and computaion are concerned there are scripts CGI, ASP and JSP, which are used in the process of analysis.

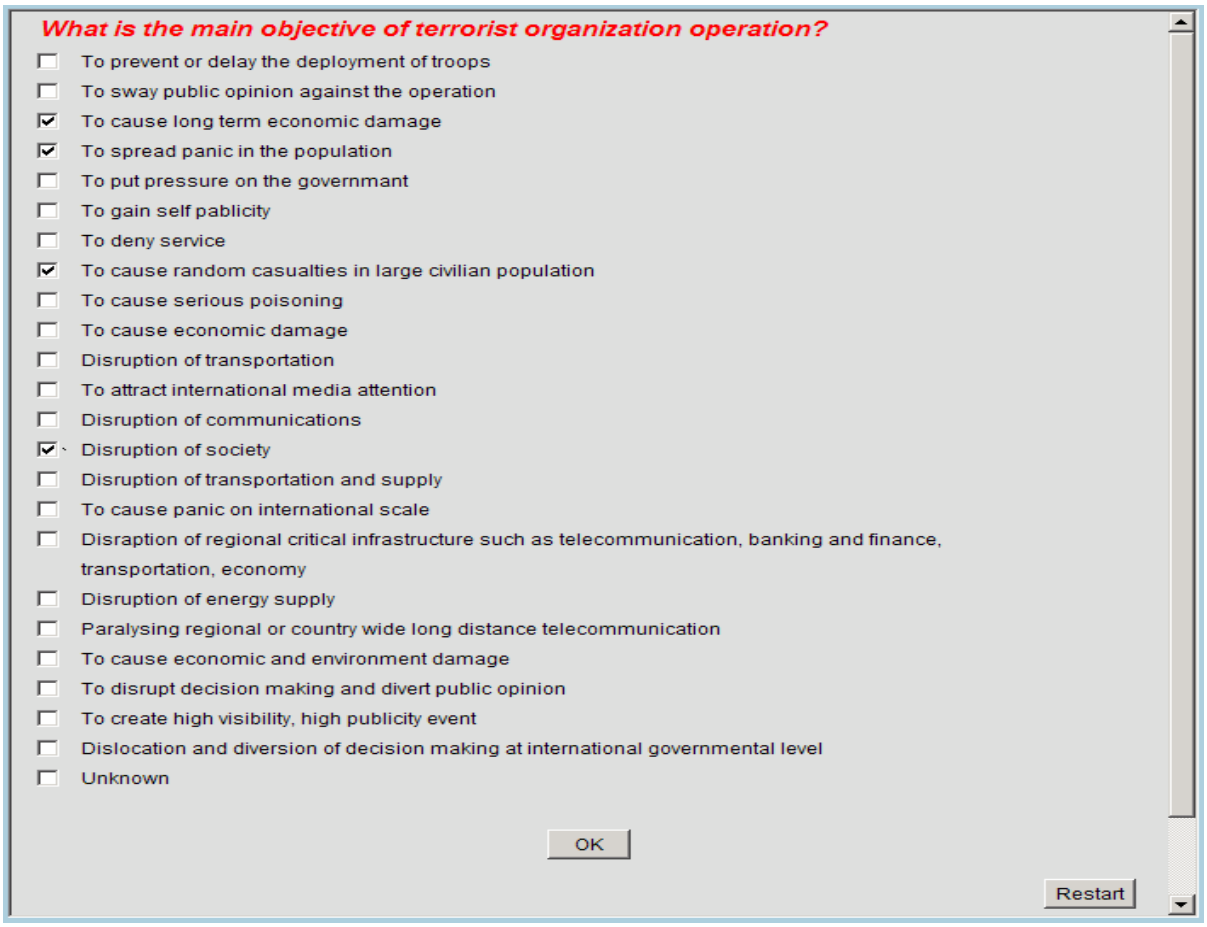


Fig.12. The prototype of expert system for terrorist identification

As the powerful environment for data mining and statistical analysis System SAS package is taken into account. Using Application Dispatcher (SAS/IntrNet) (fig.13.) the system SAS was joined to the whole environment and as remotely controlled service. System SAS is equipped with many additional tools, which can be used by users of the EWS.

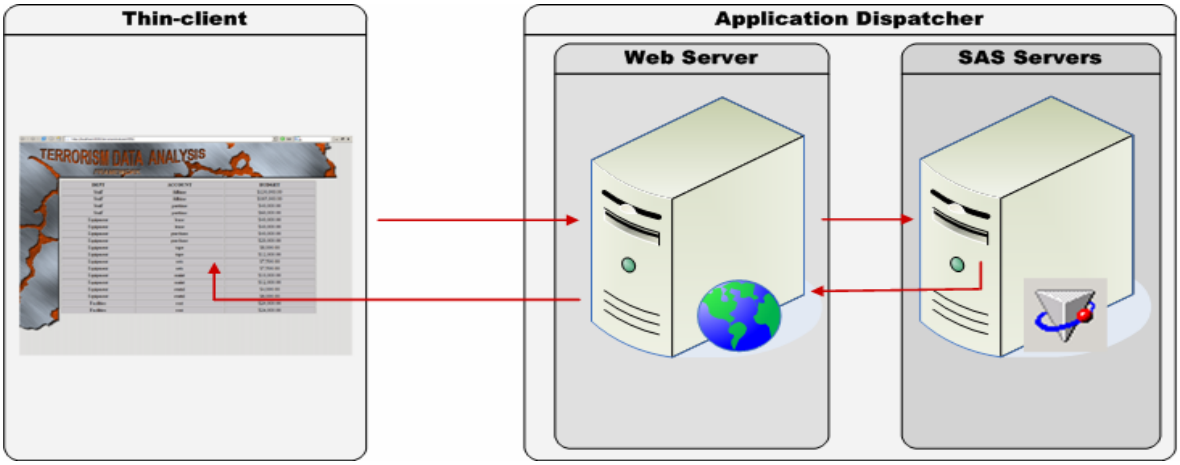


Fig. 13. The way of SAS joining,

The presented set of tools in the one environment can realize the functions of Early Warning System. These mechanisms enable the collecting, agregation and calculation of the threat coefficients finally. System SAS, which contain the neural networks builder and reggression analysis package and other data mining tools allows us the pattern recognition process.

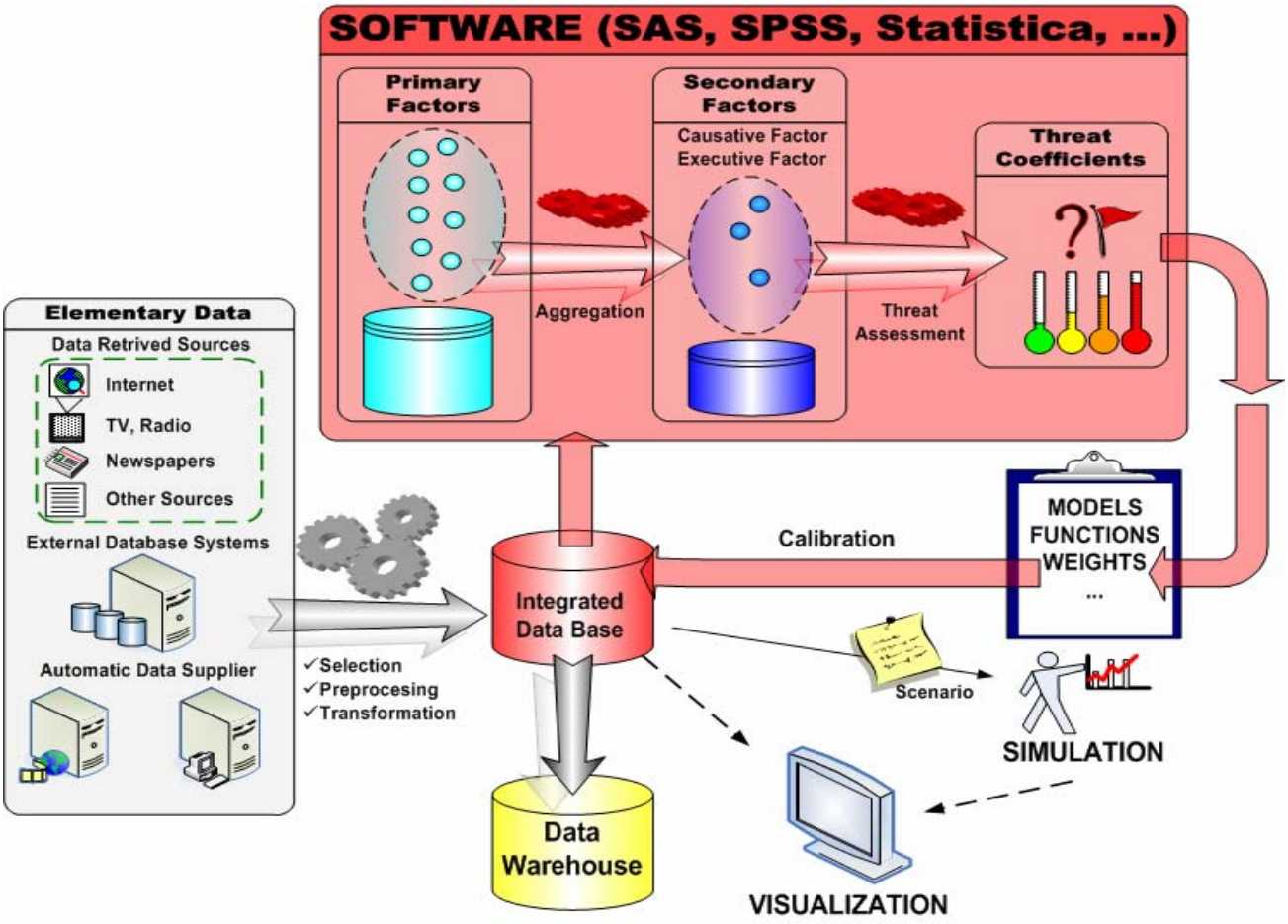


Fig. 14. The scheme of the environment of EWS

In the process of data analysis of terrorist threat we have proposed link analysis tool, which was illustrated on the basis of WTC event. The application is very useful by graphical representation of the links within terrorist organization. The tool will be developed by specific calculations of characteristics described on the net components (nodes and edges) - *clustering coefficient, average path length, shortest longest path, preferential attachment.*

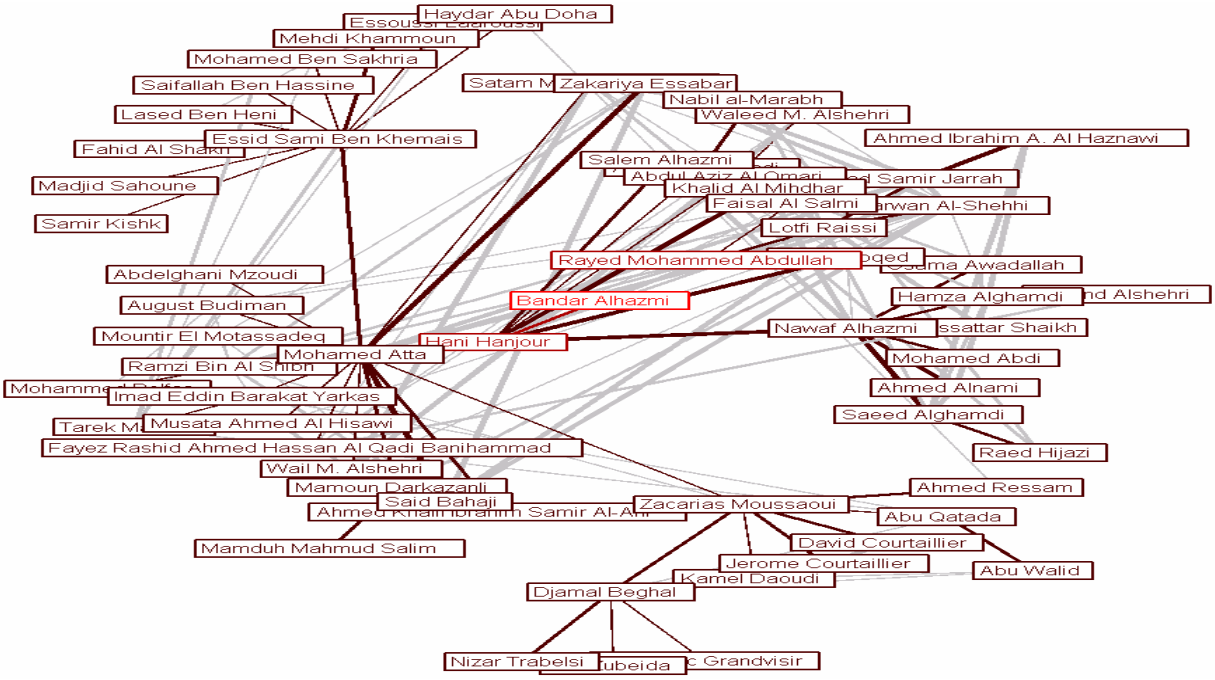


Fig.15. Representing terrorist network with application to links visualization.

The simple simulator as demonstrator of the environment possibilities was constructed and tested. The main idea there was fight between terrorist group and guard, which try to defend an infrastructure and terrain. Action is conducted on two levels:

- in the net, which is defended infrastructure,
- in the net of terrorists and guard.

The players make decisions during the game – there is interactive environment. The results of the decisions are simulated.

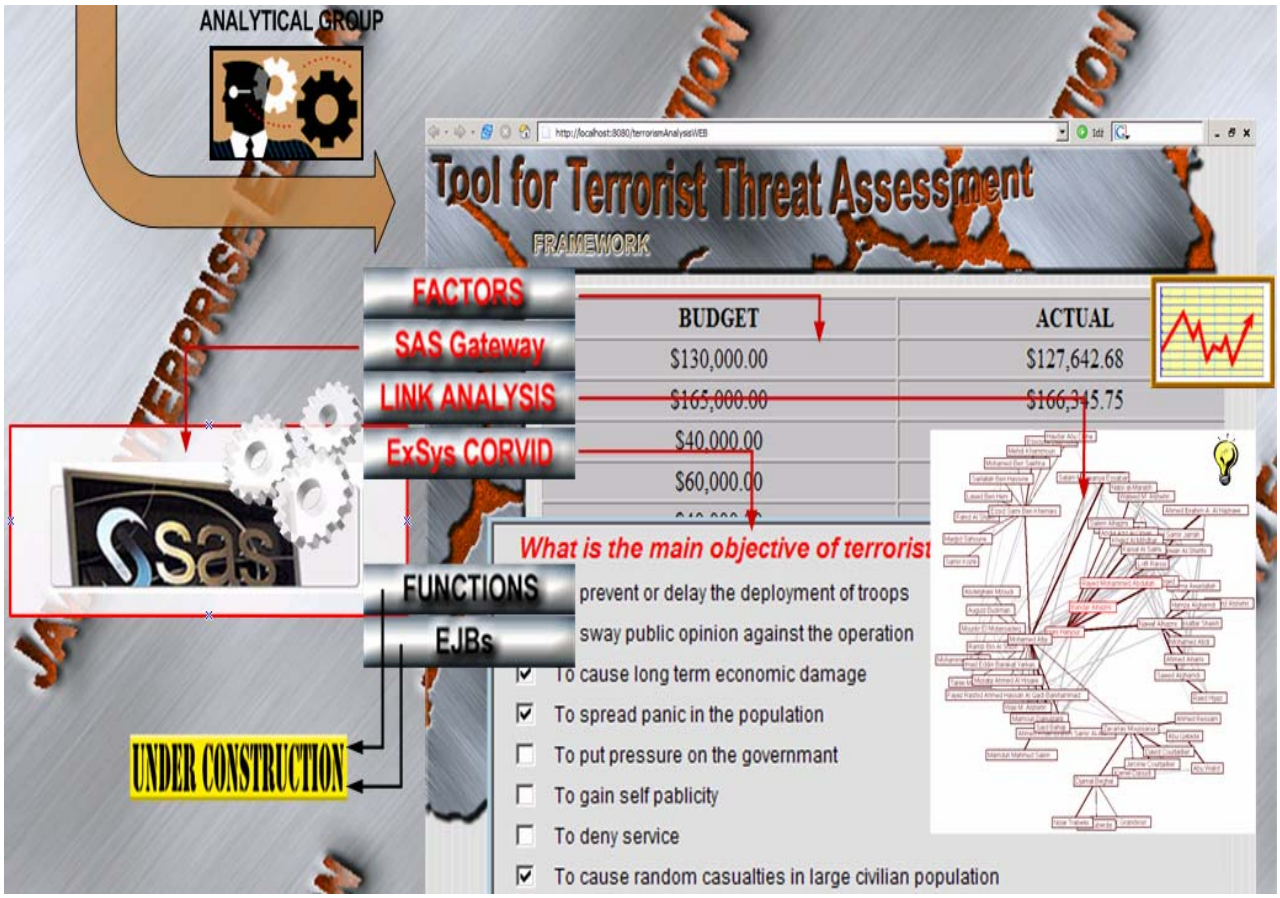


Fig. 16. The graphical user interface of the Early Warning System

7.0 RECOMMENDATIONS

Recommendations for further assessment development include that part of analysis and implementation of the tool. In the phase it is very important to have real or quasi-real set of training data (description of terrorist events).

It is believed that applying these recommendations will lead to an effective and efficient assessment framework, which not only provides the military and security operators with useful and relevant training feedback, but is a means as well for all participants to ensure that the very promising training and real threat assessment potential of EWS can indeed be achieved.

8.0 REFERENCE SECTION

[1] A.Najgebauer, R. Antkiewicz, W. Kulas, D. Pierzchała, J. Rulka, Z. Tarapata, M. Chmielewski: *A Concept of Simulation Based Diagnostic Support Tool for Terrorism Threat Awareness.* Koblenz, Germany 2004. NMSG Conference. RTO PUBLICATIONS.

[2] M. Chmielewski - *The Concept of Association Construction in Semantic Networks.* VII International Workshop for Candidates for a Doctor's Degree. VII OWD'2005.

[3] R. Kasprzyk - *Complex Networks in Countering Terrorism.* VII International Workshop for Candidates for a Doctor's Degree. VII OWD'2005.

[4] R. Kasprzyk - *Model and implementation of operational game for decision support in a class of conflict situation.* Warsaw 2005. Master's Theses under the A. Najgebauer Supervising. Military University of Technology, Faculty of Cybernetics.

[5] A. Najgebauer – *Decision Support Systems for Conflict Situations. Models, Methods and the Interactive Simulation Environment.*" Ed. Biuletyn WAT. Warsaw 1999, Poland. (294 p.). ISBN 83-908620-6-9

[6] A. Najgebauer – *Technical Activity Program of MSG 026. NMSG Publications 2003.*

[7] NATO /EAPC/PFP - *Generic Early Warning Handbook*, 2001

[8] Ripley B. D.- *Pattern Recognition and Neural Networks*, Cambridge University Press, Cambridge 1997