RTO-MP-13
AC/323(IST)TP/1

RTO-MP-13

NORTH ATLANTIC TREATY ORGANIZATION

RESEARCH AND TECHNOLOGY ORGANIZATION

BP 25, 7 RUE ANCELLE, F-92201 NEUILLY-SUR-SEINE CEDEX, FRANCE

RTO MEETING PROCEEDINGS 13

Frequency Assignment, Sharing and Conservation in Systems (Aerospace)

(l'Attribution, le partage et la conservation des fréquences pour les systèmes aéronautiques et spatiales)

Papers presented at the Information Systems Technology (IST) Symposium held in Aalborg, Denmark, 5-7 October 1998.

Published January 1999

Distribution and Availability on Back Cover

RTO-MP-13
AC/323(IST)TP/1

NORTH ATLANTIC TREATY ORGANIZATION

RESEARCH AND TECHNOLOGY ORGANIZATION

BP 25, 7 RUE ANCELLE, F-92201 NEUILLY-SUR-SEINE CEDEX, FRANCE

RTO MEETING PROCEEDINGS 13

Frequency Assignment, Sharing and Conservation in Systems (Aerospace)

(l'Attribution, le partage et la conservation des fréquences pour les systèmes aéronautiques et spatiales)

Papers presented at the Information Systems Technology (IST) Symposium held in Aalborg, Denmark, 5-7 October 1998.

The Research and Technology Organization (RTO) of NATO

RTO is the single focus in NATO for Defence Research and Technology activities. Its mission is to conduct and promote cooperative 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 coordination 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 coordinates RTO's cooperation 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 initial cooperation.

The total spectrum of R&T activities is covered by 6 Panels, dealing with:

- SAS Studies, Analysis and Simulation
- SCI Systems Concepts and Integration
- SET Sensors and Electronics Technology
- IST Information Systems Technology
- AVT Applied Vehicle Technology
- HFM Human Factors and Medicine

These Panels are made up of national representatives as well as generally recognised 'world class' scientists. The Panels 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 cooperation 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.

Printed on recycled paper

Published January 1999

ISBN 92-837-0003-1

Printed by Canada Communication Group Inc.
(A St. Joseph Corporation Company)
45 Sacré-Cœur Blvd., Hull (Québec), Canada K1A 0S7

North Atlantic Treaty Organization

Research and Technology Agency

ORGANIZATION

RTA Headquarters: 7, rue Ancelle - 92200 Neuilly-sur-Seine, France

ST/60/4 19 August, 1998

TO: Recipients of RTO Publications

FROM: Scientific Publications Executive

SUBJECT: **RTO Technical Publications**

As you probably know, NATO formed the Research and Technology Organization (RTO) on 1 January 1998, by merging the former AGARD (Advisory Group for Aerospace Research and Development) and DRG (Defence Research Group). There is a brief description of RTO on page ii of this publication.

This new organization will continue to publish high-class technical reports, as did the constituent bodies. There will be five series of publications:

AG **AGARDographs** (Advanced **G**uidance for **A**lliance **R**esearch and **D**evelopment), a successor to the former AGARD AGARDograph series of monographs, and containing material of the same long-lasting value.

MP **Meeting Proceedings**: the papers presented at non-educational meetings at which the attendance is not limited to members of RTO bodies. This will include symposia, specialists' meetings and workshops. Some of these publications will include a Technical Evaluation Report of the meeting and edited transcripts of any discussions following the presentations.

EN **Educational Notes**: the papers presented at lecture series or courses.

TR **Technical Reports**: other technical publications given a full distribution throughout the NATO nations (within any limitations due to their classification).

TM **Technical Memoranda**: other technical publications not given a full distribution, for example because they are of ephemeral value only or because the results of the study that produced them may be released only to the nations that participated in it.

The first series (AG) will continue numbering from the AGARD series of the same name, although the publications will now relate to all aspects of defence research and technology and not only aerospace as formerly. The other series will start numbering at 1, although (as in the past) the numbers may not appear consecutively because they are generally allocated about a year before the publication is expected.

All publications, like this one, will also have an 'AC/323' number printed on the cover. This is mainly for use by the NATO authorities.

Please write to me (do not telephone) if you want any further information.

G.W.Hart

Frequency Assignment, Sharing and Conservation in Systems

(RTO MP-13)

Executive Summary

Information transfer is an essential feature of modern warfare, the Frequency Assignment process is an integral part of this activity.

NATO, through the Radio Frequency Agency is responsible for the assignments in the 225 - 400 MHz band and must be conversant with developments in this field and apply the most appropriate technique. Components of NATO's forces are more frequently deployed in peace support roles where timely information is crucial and communications are in addition to the existing infrastructure. In this scenario allied forces must establish a compatible communications plan which includes frequency assignment in an already dense and dynamic environment.

Wideband systems are becoming prevalent in the civilian community and will demand greater spectral occupancy. This, coupled to the proliferation of mobile communications puts severe pressure on the Military to relinquish some of the spectrum currently allotted on an exclusive basis. Spectrum pricing will only add to this pressure. It is therefore essential that spectral resources are employed efficiently.

The symposium was organised to expose the assignment engines under development and solicit the requirements from military users. The intent was to allow academics visibility of the rules which the users wish to apply and establish an understanding of the priorities which are imposed.

The call for papers produced a good response with the subsequent presentations provoking some lively discussion. The symposium was divided into three basic groups:

- Tutorial
- Spectrum Management
- Emerging Technologies

The tutorial provided a background of the mathematical techniques which are applied allowing the other remaining presenters to focus on their particular area. A larger number of papers was received for the Emerging Technologies reflecting the activity in the civilian community to optimise the assignment engines. It is recommended that practical scenarios are made available to the academic community to exercise their algorithms in a realistic fashion for military applications. The symposium provided a suitable forum for both formal and informal discussion and gave the military community an insight into the current research areas.

G. Wyman
Programme Committee Chairman

L'attribution, le partage et la conservation des fréquences pour les systèmes aéronautiques et spatiales

(RTO MP-13)

Synthèse

L'échange de données est un élément essentiel de la guerre moderne, et l'attribution des fréquences fait partie intégrante de cette activité.

L'OTAN, par le biais du Bureau allié des fréquences radio (ARFA), est responsable des attributions dans la bande 225 – 400 MHz. Les derniers développements dans ce domaine doivent être bien connus pour appliquer la technique la plus appropriée. Les forces de l'OTAN sont très souvent employées aujourd'hui en soutien de la paix, rôle dans lequel l'actualisation du renseignement est capitale, avec des communications qui viennent s'ajouter à une infrastructure existante. Confrontées à cette situation, les forces alliées sont obligées d'élaborer un plan de communications compatible, qui permette d'attribuer des fréquences dans un environnement chargé et dynamique.

Les systèmes à large bande se rencontrent maintenant fréquemment dans le domaine civil et vont exiger de plus en plus d'occupation spectrale. Ceci, allié à la prolifération des moyens mobiles de communication, exerce une forte pression sur les militaires, qui se voient obligés d'abandonner une partie du spectre qui leur était exclusivement attribuée. L'installation d'une tarification ne fera qu'accentuer cette pression. Il est, par conséquent, essentiel d'employer les ressources spectrales de façon efficace.

Le symposium a eu pour objectif de présenter les systèmes d'attribution de fréquences en cours de développement et d'établir les demandes des utilisateurs militaires. L'objectif a été de permettre aux membres du milieu universitaire de prendre connaissance des règles que les utilisateurs veulent appliquer et de mieux comprendre les priorités.

De nombreux textes de conférence avaient été proposés et les présentations retenues ont suscité des discussions animées. Le symposium s'est scindé en trois grands groupes :

- techniques mathématiques appliquées
- gestion du spectre
- technologies naissantes

Le premier groupe a permis aux autres conférenciers de donner la priorité à leurs domaines spécifiques. Beaucoup de textes de conférence ont porté sur les technologies naissantes, ce qui est indicatif du niveau d'activité dans le domaine de l'optimisation des systèmes d'attribution dans le secteur civil. Il a été recommandé de mettre à la disposition des membres du milieu universitaire des scénarios pratiques permettant la mise en oeuvre de leurs algorithmes de façon concrète en vue d'applications militaires. Le symposium a servi de forum pour des discussions formelles et informelles, en donnant aux militaires un aperçu des domaines de recherche courants.

Click inside the blue boxes or on the titles to view the corresponding section

Contents

SESSION IIA: EMERGING TECHNOLOGY AND CRITERIA
Chairman: Dipl-Ing D. JAEGER (GE)

SESSION IIB: EMERGING TECHNOLOGY AND CRITERIA
Chairman: Mr G. WYMAN (UK)

SESSION IIC: EMERGING TECHNOLOGY AND CRITERIA
Chairman: Dr K.H. CRAIG (UK)

SESSION IID: EMERGING TECHNOLOGY AND CRITERIA
Chairman: Dr K.H. CRAIG (UK)

Theme

The successful operation of aerospace communications, radar and remote sensing systems requires effective frequency assignment. The potential sharing of the operating band with other users requires knowledge of the envelope of compatibility between the systems which are to be deployed. The requests for additional spectra from particularly the commercial sector demand that assignments are made with a view to conserving the over-subscribed frequency spectrum. The emphasis during the symposium will be on describing the techniques to optimise the packing of the requests within an allocation taking cognisance of the allotments.

The symposium will consider the frequency assignments to minimise interference and strategies which allow frequency sharing between similar and dissimilar systems. Off-line, real-time models and decision aids will be considered in the symposium together with experimental validation. Occupancy measurements and models, network planning tools and advanced computer area coverage models are important issues. Related system, networking measurements and simulations will also be highly relevant to this meeting.

Frequency spectrum conservation can be improved by operating a number of services in the same frequency band, each service employing a different modulation, coding or access scheme. An example of this is the operation of spread spectrum CDMA communications systems in a band also used by FDMA services. The symposium will address the selection of the appropriate modulation techniques and the necessary parameters for compatible operation.

In order to utilize the frequency spectrum efficiently a number of approaches have been developed around the choice of waveform. Examples of these are the high level modulation schemes (such as 16 or 64 level QAM) used in high capacity civil communications systems. The application of simple diversity techniques and signal processing to overcome the inherent limitations of the propagation channel also help in the assignment process will be addressed.

Why important to NATO: Spectrum conservation and the retention of adequate spectrum for military use will become more difficult with the proliferation of commercial wideband systems. The ability to communicate on demand is paramount for aircraft involved with safety of life issues. NATO has specific assignment responsibility in the 225-400 Mhz band in NATO Europe and advises users in other parts of the spectrum.

Thème

L'exploitation adéquate des systèmes de télécommunications aérospatiales, des systèmes radar et des systèmes de télédétection passe par l'attribution optimale des fréquences. Pour partager la bande de fréquences de fonctionnement il faut connaître l'enveloppe de compatibilité entre les systèmes à mettre en œuvre. Les demandes de spectres additionnels émanant en particulier du secteur commercial, font que les fréquences sont attribuées de façon à conserver le spectre malgré les demandes d'abonnement qui dépassent la capacité disponible. Au cours de ce symposium, l'accent sera mis sur la description de techniques permettant d'optimiser l'attribution des demandes.

Le problème de l'attribution des fréquences pour réduire au maximum les interférences sera aussi examiné, ainsi que les stratégies permettant le partage des fréquences entre systèmes semblables et dissemblables. Des modèles et des aides à la prise de décision autonomes et en temps réel seront pris en compte lors du symposium. Leur validation expérimentale sera examinée, ainsi que les calculs et les modèles d'occupation spectrale, les outils de planification de réseaux et les modèles sophistiqués de couverture informatique. Les simulations et les calculs des systèmes et des réseaux connexes seront également abordés.

La conservation du spectre de fréquences peut être améliorée par l'exploitation de plusieurs services dans la même bande, chaque service utilisant une procédure d'accès, de modulation ou de codage différente. Un exemple est l'exploitation de systèmes de télécommunications d'accès multiple à répartition par code (AMRC) à étalement de spectre dans une bande de fréquences utilisée en parallèle par des services d'accès multiple par répartition en fréquence (AMRF). Ce symposium examinera le choix de techniques appropriées de modulation, ainsi que les paramètres nécessaires à une exploitation cohérente.

En partant du choix de forme d'onde, un certain nombre de pistes ont été explorées pour utiliser le spectre de fréquences de façon efficace. Parmi celles-ci on peut citer des schémas de modulation de niveau élevé (modulation d'amplitude en quadrature QAM de niveau 16 ou 64 par exemple) utilisés dans les systèmes de télécommunications civils à grande capacité. L'application de systèmes simples de réception en diversité ou de stratégies sophistiquées de traitement du signal pour s'affranchir des restrictions des canaux de propagation et pour en faciliter l'attribution sera examinée.

La conservation du spectre est une question importante pour l'OTAN, car avec la prolifération des systèmes commerciaux à large bande, il sera de plus en plus difficile d'obtenir un spectre répondant aux besoins militaires. Et pourtant, communiquer est un besoin vital pour toute mission aérienne impliquant des vies humaines. L'OTAN est spécifiquement responsable de l'attribution des fréquences dans la bande 225 - 400 Mhz dans les pays européens de l'Alliance, et fournit des conseils aux utilisateurs en ce qui concerne le reste du spectre.

Information Systems Technology Panel

CHAIRMAN
Dr M VANT
Deputy Director General
Defence Research Establishment Ottawa
Dept of National Defence
3701 Carling Ave
OTTAWA, ON, K1A 0K2, CANADA

DEPUTY CHAIRMAN
Dr R JACQUART
ONERA/CERT/DTIM
Directeur du DTIM
BP 4025
31055 TOULOUSE CEDEX 4, FRANCE

TECHNICAL PROGRAMME COMMITTEE

CO-CHAIRMEN:	Mr G WYMAN	UK
	Mr S CAINE	US
MEMBERS:	Prof C VLOEBERGHS	BE
	Prof C GOUTELARD	FR
	Dipl-Ing D JAEGER	GE
	Prof S LECHIUTTA	IT
	Dr K H CRAIG	UK
	Dr A K HYDER	US

PANEL EXECUTIVE

From Europe:

RTA-OTAN
Lt-Col A GOUAY, FAF
IST Executive
7 Rue Ancelle
F-922OO NEUILLY SUR SEINE,
FRANCE

From the USA or CANADA:

RTA-NATO
Attention: IST Executive
PSC 116
APO AE 09777

Telephone: 33-1-5561 2280/82 - Telefax: 33-1-5561 2298/99

HOST NATION LOCAL COORDINATOR

Dr G HVEDSTRUP JENSEN
Danish Defence Research Establishment
Ryvangs Alle 1 - PO Boks 2175
DK-2100 COPENHAGEN Ø
Tel: (45) 3915 1790
FAX: (45) 3929 1533

ACKNOWLEDGEMENTS/REMERCIEMENTS

The Panel wishes to express its thanks to the Danish members to RTA for the invitation to hold this Symposium in Aalborg and for the facilities and personnel which made the Symposium possible.

Le panel tient à remercier les membres du RTB du Danemark auprès de la RTA de leur invitation à tenir cette réunion à Aalborg, ainsi que pour les installations et le personnel mis à sa disposition.

Technical Evaluation Report of the RTO Symposium Frequency Assignment , Sharing and Conservation in Aerospace Systems

Initiated by
The Sensor and Electronics Technology Panel
Sponsored by
The Information Systems Technology Panel

Donald E. Dewey
Technical Evaluation Reviewer (TER)
Mission Systems Panel Member, NATO AGARD
Boeing Company
MS 13-36
Seattle, Washington 98124, USA
and
Glyn Wyman
Technical Committee Chairman
BAeSEMA
British Aerospace
P O Box 5
Filton, Bristol BS12 7QW, UK

1. Introduction

Effective aerospace communications and sensors depends on the 1) effective management of the available (allocated) spectrum resources and on the 2) efficient use of the resources through frequency assignments. Generally, Session I covered the first topic, and Session II the second. Included in both sessions were papers from the design community of military sensors that highlighted the growing frequency assignment problem and offered potential solutions. The symposium sessions and papers are summarised and discussed in paragraph 3, below. Summary conclusions and recommendations are made in paragraph 2, below.

2.0 Summary Conclusions and Recommendations

The operation of emitters to prevent harmful interference is essential to both military and civilian users. Compatibility within the electromagnetic spectrum is achieved with a three stage process of allocation, allotment and assignment. This process is governed by the ITU, which in turn delegates some aspects of the problem to other agencies. The users maintain stability by conforming to a prescribed set of rules, lending itself to constraint logic programming. Requests for spectrum use is increasing from a wide group of both commercial and military users. Financial pressure will inevitably prevail and military users will be forced to accept smaller allocations. This symposium was organised to expose the underlying methods to achieve an acceptable position, which was accomplished. The papers offer the inclined reader a description of the method adopted coupled with an extensive bibliography.

Frequency assignments have in the past not warranted extensive research because if managers adhered to simple rules the spectrum utilisation was sufficiently low to accommodate most demands readily. The assignment process was developed from a rudimentary process of considering the separation of simple geometric shapes presented as an overlay on a map, which led to generous interference margins. This was then automated and improvements in efficiency sought by invoking sequential search processes (greedy algorithms). Current methods use one of the greedy algorithms as a primary solution and explore the solution space for improvements. Several papers were presented in this category using meta-heuristics, which were tested against standard problems as a benchmark. In some instances hybrids of the elemental methods have been described which show some promise, but all are at an early stage of development.

Methods using binary constraints predominate, but higher order constrains were covered and deemed worthy of further consideration. Linear programming methods have been considered for adoption in cellular networks and give advantages in particular areas. Regrettably the domain of application needs to be established for these techniques as well as for the hybrid methods. In this category guided local search had good performance if the internal structure can be identified, but they are competing with the neural network methods which provide considerable benefit once they have been trained on the request class.

Propagation prediction is an intrinsic element of the assignment process to enable the field strength from distant transmitters to be calculated. Several papers addressed this issue but as the efficiency of the assignment process improves, the error bound on the propagation loss must be reduced. Further refinement is required to retain consistency.

It is recommended that the theoretical methods, described in the papers presented, be exercised on practical problems to give confidence in their use. Also, the military must match the packing density achieved by their civilian counterparts to retain the spectrum currently allotted. To achieve this they must apply the latest techniques and be conversant with the techniques adopted in the civil community.

During the symposium evidence was provided which indicates that co-ordination between NATO countries is lacking. A concerted effort is required to present coherent arguments to preserve the spectrum resource; this could be achieved by establishing a working group to further the study of frequency management. Information transfer is an essential feature of modern warfare and all avenues must be kept open to ensure the greatest diversity of bearers.

3.0 Summary of the Symposium Program

Subjects for the symposium were divided into three subject groups: Tutorial, Spectrum Management and Use (Session IA & IB), and Emerging Technology and Criteria (Sessions IIA, IIB, IIC, and IID). Due to unforeseen circumstances and late insertion of a substitute paper, the papers were not categorised accurately. For clarity, the papers are discussed using the following structure:

3.1 Tutorial
3.2Sessions I and IIA: Spectrum Management and Use
3.3Sessions IIB, IIC, and IID: Emerging Technology and Criteria
- 3.3.1 Frequency Assignment Algorithms and Methods
- 3.3.2 Spectrum Needs of New Radar Concepts
- 3.3.3 Design Support Tools

Please note that for consistency, three papers given in Session I are discussed in the Session II section (paragraph 3.3). These papers are:

1) Application of New Techniques to Military Frequency Assignment
2) Real-Valued Frequency Assignment
3) Radars Basse Frequence Compatibilite Avec Les Autres Moyens Electromagnetiques (Compatibility of Low Frequency Radar with other Electromagnetic Sources)

In summarising the papers, they are at times quoted directly to maintain clarity. Many of the comments and conclusions made are those shared by both the paper authors and the TER. There are a few cases, however, where the comments and conclusions offered are that of this TER only; these can be clearly identified in the summary below.

3.1 Tutorial

The Tutorial was prepared and given by D.H. Smith and A.M. Allen from the University of Glamorgan and S. Hurley and W.J. Watkins from the University of Wales in the U.K. The intention of the Tutorial was to set the stage for those papers given in the symposium on algorithm and model research to solve the frequency assignment problem. This was effectively accomplished. The scope and seriousness of the frequency management problem to both the commercial and military agencies, and to NATO in particular, were not addressed in this Tutorial. This was accomplished later in the symposium in those papers that addressed the spectrum management and use. A summary of the Tutorial follows.

Four subjects were covered in some detail in the Tutorial. They were:

1) The two primary types of problems usually addressed in the frequency assignment process, "minimum span" and "fixed spectrum"
2) A few of the more significant algorithmic approaches; most, but not all of the algorithms covered by the papers given at the symposium were discussed in varying amounts of detail.
3) The process for evaluating algorithms using lower bounds, assuming a set of binary constraints
4) The general assumption that binary constraints is the best way to represent the problems in the real world was questioned, and the value of considering other approaches was discussed.

The minimum span problem is where it is desired to determine the minimum spectral requirement for a given service, whilst not exceeding a defined maximum level of interference. This problem receives the greatest attention in literature. The fixed spectrum problem is where it is desired to minimise some measure of interference using only a given allocation of frequencies. This problem is one that concerns frequency managers who assign frequencies within a given frequency allocation.

The Tutorial included a survey of the following algorithms:

- Exact - applicable to the minimum span problem; however, the algorithm usually finds solutions for very small problems (less than 30 transmitters)
- Partial (or limited iterations of the algorithm) Backtracking - applicable to minimum span problem, but unlikely that it will have as much flexibility in exploring the search space as the best heuristic algorithms
- Sequential (or greedy) - useful for minimum span problems, but are used to assign frequencies to transmitters with no reassignment possible
- Neural Networks - described as not being competitive in value to metaheuristic algorithms for finding optimum frequency assignments

- Genetic - an algorithm that simulates some of the processes of evolution and natural selection, and is being used for minimum span problems
- Hill Climbing - can be used to find moderately good solutions to the fixed frequency problems
- Simulated Annealing (SA) - a metaheuristic algorithm derived from statistical mechanics; can be adapted to minimum span problems
- Tabu Search - same as SA, above

Evaluating these algorithms, when applied to minimum span problems, is usually accomplished in context with a set of binary constraints. Once the constraints have been established, then the question is "how good is a particular algorithm at generating optimal solutions given such problem definitions"? Establishing how good a solution is judged to be is done by comparing how well it performs to solutions obtained by other algorithms. To do this, it is very important that common problems and their datasets be used. Smith and Allen suggests that either the CELAR dataset or the "Philadelphia" problem be used.

When comparing a solution made by an algorithm for a minimum span problem, it is necessary to establish a lower bound. This means that for a solution to be considered acceptable, it must have a solution that is within a specified range of the best possible solution found by any algorithm so far. For commercial cellular communications this value has be set at 0 - 5%. For military radio communications, the value is 0 - 10%.

Evaluating algorithms when applied to fixed spectrum problems is not as simple as for the minimum span problem, because it is not obvious how a lower bound can be established for comparison. D. H. Smith referred to some previous work done on this problem, and is referenced in the tutorial literature.

The Tutorial included the results of a number of algorithm solutions using the method described above. The metaheuristic algorithms appear to consistently provide the best or "optimum" solutions.

Although algorithm evaluation usually assumes a set of binary constraints, Smith asserts this might not be an assumption that will produce an optimum solution. Binary constraints are usually limited to describing the necessary channel separations to sufficiently attenuate interfering signal. This approach assumes there is only one possible interferer, which is not always a situation found in the real world. Another approach was presented that assumes a non-binary situation, and the non-binary and binary methods were compared. The results of the comparison suggests that non-binary constraints have a role in the evaluation process, but perhaps these constraints should be inserted into the process as revisions to the original constraints rather than using them in some way as initial conditions.

This is an excellent Tutorial for those who wish to learn more about the comparative value of various algorithms being used today for finding optimum solutions to the frequency assignment problem.

3.2 Sessions I & IIA: Spectrum Management and Use

Two papers in this session described the spectrum certification process (S.R. Green and C.A. Scammon) and the spectrum management system and tools (T.C. Hensler) being used in the US. Until recently, spectrum certification process in the United States Department of Defense (DoD) used the DD 1494 application form to collect the technical information to permit officials to evaluate the request for use of electronic spectrum frequencies. Green and Scammon observe, however, that the data collected by this process was rarely correct. The form itself has been confusing, and with the emergence of more sophisticated, complex systems, the form did not provide a means to enter sufficient details about the systems' operating characteristics. They estimate that the application could exceed 30 pages of technical information, and because the size of the form is inherently limited, the information necessary to evaluate the application could be truncated or even deleted.

In the 1990's, the Joint Spectrum Centre developed an automated system, called the Spectrum Certification System (SCS), designed to improve and enhance the Spectrum Certification process. The DD 1494 process was automated, and data checking was added to make sure the data was accurate. However, the process is not without problems. For example, if a system has several power levels and different emission bandwidths, NTIA needs to know that power levels are associated with which emission bandwidths to perform an accurate EMC analysis. Without this input, the EMC analysis performed is very conservative, and a "worst case" analysis is created.

To solve the problem of inadequate data, the US National Telecommunication and Information Administration (NTIA) is again redesigning the spectrum certification process and also is developing a new system for capturing and reporting the needed technical data. This new system, called the Smart Interface Diagram (SID), provides an icon-based interface that captures inter- and intra-system relationships and that prompts the applicant to enter only those system parameters required. This new system is described by Green and Scammon in their paper.

T. Hensler described the United States Joint Spectrum Management System (JSMSw), how it is mechanised on a set of management tools that run on a personal computer, and the eight steps that must be performed during the frequency nomination process. Recently the NTIA joined forces with the DoD Joint Spectrum Centre (JSC), and the JSMSw will be used by both agencies for commercial sector and DoD applications. There are approximately 700 copies of JSMSw being distributed to DoD and Federal users in the US. It appears that JSMSw can be used at the NATO level. This possibility should be explored in the future. The major drawback of the program is that it is used only in certain frequency ranges, and while presently being used in ground to ground applications, it has not been applied to ground-air-ground and air to air applications

The needs for spectral efficiencies for NATO were discussed by A.W. Graham and J.B. Berry from the UK. They agree with the concept being pursued by the US that a computer-based spectrum management program can play in the efficient, optimal design of networks and frequency utilisation for tactical applications (where frequency assignments need to be made dynamically), but they add that there is an urgent need for the efficient use of the military frequency band allocations. They point out that there is not only a growing competition for frequency space between the civil and military users, but there is growing competition within the military itself. For example, there is a growing needs for more frequency space to support command and control communications particularly to accommodate the digitisation of the battlefield concepts; and also to introduce battlefield surveillance radars discussed in 3.3.2. Unfortunately, NATO member nations still tend to manage frequency assignments and overall frequency allocations by paper, with varying degrees of computer support. Even with some support from computer-based systems, frequency assignments are not optimum, and are often difficult to implement preventing the manager reacting quickly to dynamic scenarios. Many computerised frequency assignment systems make it possible to improve the accuracy of the data, and can optimise frequency use by using an algorithm. However, the authors point out that although there are likely to be improvements in use, the key issues governing spectrum "lies elsewhere". The main factor for a allowing a spectrum manager to meet the requirements of an assignment in a short notice is they must have adequate, accurate, and timely information on current allocations and assignments for both military and civil agencies, and information on the environment.

The authors identify three main information requirements. The first is the need for an equipment database that includes friendly, enemy and neutral emitters and receivers. This kind of a data base does not exist (nor does a strategy for capturing a database exist), and it will require much effort to form this database. The second main piece of information needed is the dynamic order of battle (ORBAT), which includes information about friendly, enemy and neutral assets. Unfortunately the ORBAT database does not normally include all of the information required, and other sources of information is required. The third kind of information required is that of environmental data. Some general data is usually available to the manger, but since atmospheric conditions are dynamic and complex, it is unlikely that a real-time, theatre-wide picture of small scale variations in the atmosphere can be created practically for the manager in the near future.

K. Kho and M. Elliott from the NATO HQ C3 Staff described the NUFAS 2 assignment system (a software tool) being used by NATO to manage the Air-Ground-Air (A/G/A) assignments in the UHF 225-400 MHz military band. It is designed to assign frequencies for both bulk assignment problems (when it is necessary to reorganise UHF bands) and individual and batch assignment problems (for day to day assignments).

This paper is timely and enlightening. However, it is distressing to note (as discussed in the summary) that some of the work going on in the individual NATO nations are not being co-ordinated adequately with the NATO HQ C3 Staff.

Two software planning aids used in the UK to perform simple tasks, such as estimating usable frequencies, and complex tasks, such as frequency allocation for large networks, were described by N. Wheadon. These two planning aids do not replace the Automatic Link Establishment (ALE) and the system Automatic Link Maintenance (ALM) systems. It was emphasised that ALE and ALM demand a good basis, this is achieved by a Frequency Management tool. The techniques are thus regarded as complementary.

The first software planning tool, WinHF, is a combined ground-wave and sky-wave propagation prediction program developed by GEC-Marconi. The ground-wave prediction algorithm used in WinHF applies the Bremmer method, and the sky-wave element is provided by the "ITU-R" method. WinHF is used by all three services in the UK MoD, and is their de facto standard at the present time. The second software planning tool is also used across all three services in the UK MoD, and is used for generating and evaluating frequency assignments for very large systems (more than 1000 individual networks, each consisting of many radios). The primary aim of the tool is to minimise interference. This very sophisticated tool provides an output in the form of lists of frequencies for each of the networks, as well as descriptions of the quality of the frequency assignment.

This excellent paper was accompanied by a demonstration of the software tools. These tools should be reviewed by other NATO countries for possible use.

ELMER S.p.A is developing a flexible, reliable HF radio communication system for use between moving, ground-based operating platforms and ground-based, fixed stations. As a part of this development program methods were developed to minimise the number of ground stations required to achieve reliable coverage. M. Proia and G. Maviglia presented a paper on the methods used, which employ the judicious selection of ground-based station locations, knowledge of the variable conditions of the ionosphere layer at specific frequencies (based on observations made by the National Institute of Geophysics in Rome, Italy), and use of a software simulation program (ASAPS) to establish reliable HF communications with the system. The method has been demonstrated in tests in the Mediterranean area. Proia and Maviglia suggested at the symposium that this work should be of benefit to NATO for the development of reliable HF communications. Follow-on briefings to NATO are planned.

Other work going on to develop design processes and supporting tools for designing and optimising communication networks is being funded by the EU ESPRIT Project 23243. This work, described by S. Hurley, S. Chapman, and R. Kapp-Rawnsley, is being developed primarily for designing and optimising commercial mobile (cellular) communication system networks. The models and algorithms developed in this program have provisions for considering optimum frequency assignments (for reliability and non-interference), site location, and cost. Since site locations for commercial cellular systems are usually limited (due to prohibitive land costs or other constraints) or are sometimes already existing, the model used for optimisation includes features that will enable the designer to start with no initial sites (the authors refer to this condition as the Greenfield process) or with pre-existing sites (called the expansion process). Once the design process is started from the initial requirements and constraints, a number of trial networks are generated. Changes in the networks usually are made at the site configuration, where three types of antennas (omni-directional, narrow panel, and large panel) are used, and in the frequency assignments. A simulated annealing algorithm is used to determine if an improvement is made over the previous trial network. If so, then the network is modified for improvement, and evaluated again. If not, then the previous network is used to generate a new trial configuration.

Although the work described in this paper is impressive, it is focused entirely on the commercial cellular communications problem. While not directly applicable to military functions, much of this work can be used, with modification, to military and NATO tactical problems.

3.3 Sessions IIB, IIC, and IID: Emerging Technology and Criteria

Solutions to the frequency assignment problem (FAP) is usually stated as being one that minimises the spectrum requirement by finding a satisfactory assignment while using the smallest number of distinct frequencies. A second, and sometimes an equally or even greater important requirement is the solution quality should be maximised by finding an assignment using a given fixed allocation of frequencies that provides the best possible coverage. This is usually the case for commercial cellular communications, and likely for NATO C3 as well. The problem is, then, by using only simple binary constraints often assumed in many FAP solution methods, gaps in the service area occur. This point was made by Smith in the Tutorial, and is the point made by J. Bater, P. Jeavons, D. Cohen and N. Dunkin in their paper which asked the question "are there effective binary frequency separation constraints for the frequency assignment coverage problem"? Using the software package FASoft and the Pinkville FAP model, the study showed that the simple binary constraints produced a good solution in terms of frequency allocation and span. However, the coverage was reasonable, but not "100%". "Prudence" factors were added to the binary constraints until 100% coverage was achieved, which meant that to guarantee complete coverage for this Pinkville FAP, an additional 17dB was needed in the binary constraints (from 9dB to 27dB). Of course this had an adverse effect on the frequency span. The authors then selected a global constraint value that was more "sensible" rather than the "extreme". This resulted in an improved frequency span solution, achieved a good solution to the frequency assignment problem with 95% coverage. So, they proved in this case that a global constraint resulted in a better solution for their FAP than either the straight binary or modified binary constraint.

Not found in the Bater, Jeavons, Cohen, and Dunkin paper, unfortunately, was a presentation at the symposium about some recent work they have done on the development of a "Higher Order" set of constraints; higher order constraints as they defined them looks at multiple, rather than single (or binary) interferences. These constraints have permitted them to obtain solutions (still using FASoft) that are much more optimum than when binary constraints, prudent binary constraints, or global constraint approaches are used. This work described by Bater, Jeavons, Cohen, and Dunkin is encouraging. Follow-on work is needed to determine how well the higher order set of constraints, as proposed, can be adapted to function and perform with other algorithms.

R. Bradbeer presented a paper on the application of simulated annealing (SA) and genetic (GA) algorithms to the problem of efficient and optimal frequency assignment problem, but with a major improvement. As Bradbeer points out, the two algorithms can not be simply and directly applied to FAP, because the problems to be solved (genetic or annealing vs. frequency assignment) are not fully analogous. In real annealing, for example, the atoms (which are identical and the constraints governing their relative positions are the same) are moved in small increments until they occupy the correct position in relation to their neighbours. In the frequency assignment problem, the constraints (i.e. cosite vs farsite) for each frequency application are not identical, and the individual frequencies are interrelated in a complex relationship. Therefore, simple application of the simulated annealing algorithm, which does not take into account the relative interaction, will not result in an efficient or even desirable solution.

Simple application of genetic algorithms is flawed as well. An application of this algorithm might involve the generation of frequency assignments where there is the lowest overall interference. At this point, however, there easily could be those frequency assignments that would interfere with one another, so there would have to be further work to eliminate the interference. The genetic algorithm assumes (appropriately) that crosses between entirely different organisms are not viable, so this algorithm is not appropriate for this step. The need for speciation in genetic algorithms is addressed in the literature, but the total solution space required for assigning hundreds of frequencies to hundreds of nets is enormous, and the author has not found any reference to solving this kind of problem.

Bradbeer proposes that if the simulated annealing or the genetic algorithms are used and practical results are to be achieved, then productive areas of the solution space will need to be identified and the algorithm "guided" so as to avoid exploration of unproductive areas. Alternatively the techniques can be applied in such a way that the problem presented to the algorithm has a smaller solution space and is less multi-modal. An example of this latter technique is proposed by Hurley et al and is referenced in his paper. Bradbeer chose to guide the algorithm using the standard graph colouring algorithm developed by M. A. Trick from Carnegie Mellon University. Although the results of this approach are not covered in his paper, the approach deserves to be seriously considered by others for future application.

It is well known that some algorithms are better at satisfying the constraints, and others are better at finding better cost solutions. Following this logic, C. Bousano-Calzon and A. Figueiras-Vidal have developed a hybrid algorithm that merges the best features of the neural network algorithm with the best features of the genetic algorithm.

The genetic algorithm used by the authors uses the same approach given by Goldberg. The neural network used is a Hopfield type net similar to that in Funabiki, but with some new properties to make it more compatible with the genetic algorithm. Quoting the authors in their paper, "to mix the neural net and the genetic algorithm, each gene represents and initial point in the neural net. The fitness of the gene is then evaluated by first initialising the neural net, relaxing it to a feasible solution and counting up the assigned frequencies or channels".

The approach taken to evaluate the hybrid algorithm was to (a) analyse the performance of the neural network alone on two FAP's and then to (b) compare the performance of the hybrid algorithm to the neural network performance and the performance of a genetic algorithm alone as reported in the literature. The authors selected a general form of allocation by Gamst and Rave for arbitrary non-homogeneous networks. The results of this study showed that the neural network portion of the hybrid algorithm allowed the algorithm to have feasible solutions in each case. The quality of the allocations obtained by the network was approximately 2% below the optimum for the benchmark problem.

Studies into the hybrid algorithm (such as the one developed by the authors) should continue. The basic concern here, however, is how the performance was measured and studied. The test case problem is much too generic; as pointed out in the Tutorial, standard problem cases should be used in studies of this type.

C. Voudouris and E. Tsang also use a derivation of the neural network algorithm to work with local search algorithms to solve the Radio Link FAP (RLFAP). Local search algorithms alone have not led to optimum solutions for large problems, but they are very effective for small problems. Voudouris and Tsang use this new algorithm, which they derived from the GENET neural network and which they call Guided Local Search (GLS), to "guide" the local search toward optimal solutions for the total RLFAP.

In their paper, they show how well the GLS and variants of the GLS solved the RLFAP relative to publicly available performance benchmarks. For the 25 different instances of the standard RLFAP, the GLS achieved a better solution in four instances, did marginally poorer in two instances, and achieved the same results in all the rest. The variants of the GLS achieved very good results as well. In the thousands of runs made testing the variants, they were all able to find a feasible solution of the soluble instances despite the fact hard inequality constraints were included in the cost function. When a circular list strategy was used with the GLS variants, better and more consistent solutions were found, compared to the best known solutions found by all the different algorithms in use today. This leads the authors of this paper to make the bold statement that "we believe there is little room for improvement with regard to solution quality by using more sophisticated techniques."

The authors have good reasons to be optimistic about the GLS, and particular its use when circular list strategy is used. One very major step must be taken first, and they stated so in their presentation and in the last sentence of the paper. That step is to see how well the GLS performs in the real world, working on real FAP problems, and how easily the GLS can be integrated into the frequency management problem. This is a next step that is highly recommended by this TER.

R. Leese reported on the results of an effort to refine the linear programming approach for FAP the frequency assignment problem. The refinements include utilising linear programming relaxation, and adding column generation to solve the problem. This research was funded by the UK Radiocommunications Agency under a contract with St. Catherine's College, Oxford.

Linear programming has produced promising results for a variety of standard benchmarks. The main new refinement proposed by Leese is the idea of adding column generation, which is called the Channel Assignment by Column

Generation (CACG) procedure. The CACG procedure has been applied to the range of problems discussed in the Tutorial paper by Smith (see Tutorial paper, above), and results have been reported by Leese in the Proceedings of the 14gh International Wroclaw Symposium of Electromagnetic Compatibility, June 23-25, 2998.

The CACG procedure is a combination of several standard techniques from linear programming optimisation. As a first step, CACG "relaxes" the requirement in the linear programming (the simplex method is used) that all variable be integer. This 'relaxed' linear program makes it much easier to solve a problem, but it produce a fractional solution that is not possible to occur. To overcome this problem, rounding is used in CACG. Column generation is applied to the relaxed linear program, where the demand constraints (ie. sufficient number of channels) are introduced. The process taken here involves the use of building blocks. An 'auxiliary problem' is used to construct the building blocks. Interference constraints are introduced through the auxiliary problem.

Leese shows that CACG has the potential to be a good method whenever there is an optimal assignment that may be constructed from spectral blocks. The CACG benchmarks show good performance for many instances, and also show there is a need for improvements in other certain instances. Leese observes that further research is needed to clarify the range of applications of this type of method.

3.3.2 Spectrum Needs of New Radar Concepts

As pointed out before, there is an ever growing demand for frequency space from both commercial and military users. New low observable radars, for example, need to operate in the V/UHF bands where they will be in competition with other radio communication devices. In their paper J. Zolesio and B. Olivier describe a proposed low frequency radar for detecting aircraft vehicles such as helicopters. The reason for operating in this lower region (200 to 500Mhz) is to improve the detection of slow moving aircraft and missiles. However, this region is heavily used for communication, requires larger antennas, and is usually easy to detected. Zolesio and Olivier propose four strategies for overcoming these problems. First, they propose to use a 20db gain antenna. This kind of antenna is very directional, reducing the spatial occupancy and hence will reduce the chances of it being detected and should reduce interference with other friendly assets. Second, the transmitted frequency is heavily filtered, with sidelobes below 40db. This also reduces the probability of detection and interference with other devices operating in the V/UHF region. Third, they propose to limit the total transmit power to 100 watts and limit the pulse to short bursts. Fourth and last, they are proposing a frequency hopping strategy over a very wide band of frequencies, which will again reduce the probability of detection. However, it was on this point that their concept posed a problem in frequency assignments. The frequency hopping concept uses many frequencies, and is quite unpredictable.

Obviously there is a problem for sharing the V/UHF space with other assets. Although long distance surveillance radars operating at these frequencies operate using kilowatts of power, mapping radars looking for ground targets between 10 and 15 km use about 100 watts. Therefore, these kinds of mapping radar should not be likely to interfere with other assets operating at the V/UHF frequencies.

A new type of radar concept that need large portions of the frequency band is the wide band and the ultra-wide band radar. Jean Isnard, in his paper, explains that while present radars use about 1% or so of the available band width of the carrier frequency, wide band radars can use as much as 10%, and the new experimental ultra-wide band radars will use 25% of the carrier. So if these new radars require so much, why are they necessary? Isnard gives the following reasons:

a) they can achieve much higher resolution
b) lower probability of detection (low observable) of the radar by hostile forces because the radar energy is spread over the spectrum
c) in the search mode, clutter is reduced
d) in the terrain following mode, multipath effects can be eliminated
e) in the recognition mode fratricide fire can be eliminated

Although ultra-wide band radar will use as much as 25% of the carrier frequency, they also operate in short bursts (on the order of one nanosecond); most of the time they are just listening. So when the time-frequency space required for ultra wide band radar is compared to other types of radio-communication assets, the problem is not as great as one might assume. However, it is obvious that wide band radar designers and frequency spectrum managers need to have much more interaction with the communications community, as Isnard states. Designers have options and design techniques that may help alleviate much of the problem. For example, ultra wide radars are being designed with very wide band receivers, which have innovative signal processing to eliminate other interfering electromagnetic transmitting sources. The technology developed could have application to other designs. If the receivers are tolerant to interference then a greater packing density can be achieved and the Frequency Manager can relax the rules applied.

3.3.3 Design Support Tools

The growing number of commercial and military users of the electromagnetic spectrum has placed pressure on the system designers to consider using the higher millimeter frequencies, where the spectrum in relatively uncongested and wide bandwidths are more readily available. In their paper, A. Shukla, A. Akram, T. Konefal, and P. Watson describe a decision aid tool, called the Millimeter Decision Aid System (MIDAS) for use by designers who are developing hardware operating in the 20-40 GHz range, and for system planners to match communication tactics to the battlespace environment to maximize the operational effectiveness of the communication assets. MIDAS includes four models: tropospheric, propagation prediction, communications prediction, and system recommendations.. It also includes an input/update module for human operator interaction with the four models. An output module is included to display the predictions to the operator in an effective manner. The tropospheric model contains the following submodels:

- rain - method used is that of Tattelman and Scharr
- melting layer - which is a mixture of ice-water and air, is treated as rain that is assumed to extend to 0 degrees C
- cloud - derived from the Salonen and Uppala model
- water vapour - uses a combination of the Salonen and Uppala annual water vapour content and surface value of water vapour content equations
- oxygen - uses an atmosphere model

The propagation and attenuation model contains the following submodels:

- rain attenuation - rainfall rates are converted into attenuation using the method of Leitao and Watson.
- cloud attenuation - uses Salonen and Uppala
- water vapour attenuation - uses modified CCIR Radiometeorological data
- oxygen attenuation - modified to account for rain, cloud, and water vapour attenuation

The present MIDAS configuration has three primary outputs: 1) signal coverage maps of either signal availability for a fixed excess attenuation or excess attenuation for a fixed availability; 2) colour contours of signal availability and link margin that need to be adjusted; 3) point to point link budget analysis.

Even in it's present configuration, MIDAS is a useful tool for designers and planners. It appears that it could be a useful tool for NATO operations planners, since it has been developed to cover most of Europe. However, it does not appear that MIDAS is very close to being ready for use as a tactical, battle field planning tool, since there were no military constraints considered in the model. Perhaps after the planned models are added in the future this may change.

One last paper presented by J. Bohl, T. Ehlen, and F. Sonnemann covered the subject of high power microwave (HPM) effects on LF electronic circuits. Whilst important, it was of peripheral interest to the participants of this symposium. There was some value in having the paper given at the symposium, however; those who are concerned about frequency allocations, assignments, and overall management need to be reminded often of the potential detrimental effects electromagnetic radiation can have on sensitive, friendly assets in the battlefield.

One particularly interesting point made by the Bohl et al is that HF radiation can be converted at the interface of some electronic devices to LF interference. So, while most electronic devices can be and are usually shielded from HF energy, the devices are often susceptible to LF interference generated at the device interface. Therefore, simple shielding and other isolation techniques may not protect electronic devices.

The authors review the work they have done to develop appropriate simulation tools for analysing and calculating the interference using numerical techniques. Network simulation tools are essential, but only for looking for interference effects. In real life, the interference effect may be some disruption of the function of the electronic device, and in real cases where the HPM is very high, the device may even be destroyed. At the present time, network tools can not be used for this purpose. Real measurements must be made to do this; the authors discuss how they use a generic device that models the actual electronic device for this kind of analysis.

For those interested in this phenomenon, this paper should be worthwhile reviewing.

Frequency Assignment: Methods and Algorithms

D.H. Smith and S.M. Allen
Div. Maths. and Computing
University of Glamorgan
Pontypridd
Mid Glamorgan CF37 1DL
Wales, U.K.
email:{smallen, dhsmith}@glam.ac.uk

S. Hurley and W.J. Watkins
Dept. of Computer Science
University of Wales, Cardiff
P.O. Box 916
Cardiff CF2 3XF
Wales, U.K.
email:{s.hurley, w.j.watkins}@cs.cf.ac.uk

1 Summary

In this tutorial paper the representation of the frequency assignment problem as a generalised graph colouring problem is presented. Both minimum span and fixed spectrum problems are described. A number of algorithmic approaches are outlined, with particular emphasis given to simulated annealing and tabu search. Hybrids of these algorithms with sequential algorithms and approaches based on subgraphs are also mentioned. The use of lower bounds to evaluate algorithms and assignments is explained. Finally, consideration is given to the effectiveness of the generalised graph colouring approach. An alternative approach using non-binary constraints is described and a method of generating lower bounds for this non-binary model is introduced.

2 Introduction

The frequency assignment problem arises when a large number of transmitters are operating in a region, and the interference caused by transmitters on receivers served by other transmitters is to be minimised. This can normally be achieved easily if the number of available frequencies is large enough. The problem is to minimise the interference while at the same time using spectrum efficiently. The frequency assignment problem arises in many military applications, such as networks of point to point radio links, and in civil applications such as cellular mobile telephone networks.

There are two essential types of problem that can be tackled. The first problem, known as the minimum span problem, attempts to determine the minimum spectral requirement for a given service, while maintaining a defined maximum level of interference. The second is known as the fixed spectrum problem, and attempts to minimise some measure of interference using only a given allocation of frequencies. The first problem has received the greatest attention in the literature, whereas the second is the more important for the majority of frequency planners, who have to operate within some given frequency allocation.

Frequency assignment problems can be represented as generalised graph colouring problems. The use of this representation goes back at least to 1970. Zoellner and Beall [1] and Lanfear [2] both cite Metzger [3] as the earliest author to use this approach. Zoellner and Beall regarded the approach as a breakthrough in frequency assignment technology. More recently this approach has been questioned. The evidence for and against the generalised graph colouring formulation will be discussed later in this paper.

In the generalised graph colouring formulation a constraint matrix $[\phi_{ij}]$ is defined such that if f_i denotes the frequency assigned to transmitter i, then in order to avoid interference it is required that $|f_i - f_j| > \phi_{ij}$. As these constraints each involve exactly two transmitters they will be referred to here as binary constraints. Sometimes there are constraints concerned with intermodulation products. These are not binary, as they involve more than two transmitters. For example, they might have the form:

$$2f_i - f_j \neq f_k$$
$$\text{or}\ \ 3f_i - 2f_j \neq f_k$$

Consideration of non-binary constraints will be deferred until section 5. There can also be constraints which imply that the frequencies assigned to two specified transmitters differ by exactly some fixed number of channels. When these occur they usually represent the two directions of a fixed link. The existence of such constraints can be used to reduce the size of the problem under consideration. They have been considered particularly with respect to problems related to

Paper presented at the RTO IST Symposium on "Frequency Assignment, Sharing and Conservation in Systems (Aerospace)", held in Aalborg, Denmark, 5-7 October 1998, and published in RTO MP-13.

the CELAR dataset used in the CALMA project[3]. Such constraints will not be considered further here.

The frequency assignment problem can be represented in a graph theoretic formulation. The transmitters are represented by the vertices $V(G)$ of a constraint graph G [4]. Vertices are joined by an edge if and only if there is a constraint between the frequencies assigned to the two transmitters they represent:

Definition 1 *A* constraint graph G *is a finite, simple, undirected graph in which each edge* $v_i v_j$ $(v_i, v_j \in V(G))$ *has a non-negative integer label* ϕ_{ij}.

A constraint graph can be seen in Fig. 1 and the associated constraint matrix is shown in Fig. 2. As the constraints are symmetric, the matrix is shown in upper triangular form to avoid repetition, and 'n' denotes 'no constraint'.

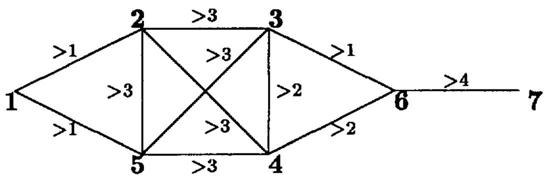


Figure 1: An Example Constraint Graph

.	1	n	n	1	n	n
.	.	3	3	3	n	n
.	.	.	2	3	1	n
.	.	.	.	3	2	n
.	.	.	.	.	n	n
.	.	.	.	.	.	4
.	.	.	.	.	.	.

Figure 2: The Constraint Matrix of the Graph in Figure 1

The vertex colouring problem for graphs is the problem of finding the minimum number of colours necessary for colouring the vertices of a graph in such a way that adjacent vertices are assigned different colours. This corresponds to minimising the number of frequencies necessary when the only constraints are cochannel constraints, which are binary constraints of the form

$$|f_i - f_j| > 0.$$

An account of many colouring problems in graphs can be found in [5]. As the frequency assignment problem involves binary constraints other than cochannel constraints, it has to be regarded as a generalised form of the graph colouring problem.

Definition 2 *A* frequency assignment *(or channel assignment) in a constraint graph G is a mapping $f : V(G) \to F$ (where F is a set of consecutive integers $0, \ldots, K$ representing frequencies) such that the constraints*

$$|f(v_i) - f(v_j)| > \phi_{ij}$$

are satisfied for all $v_i v_j \in E(G)$. Sometimes this is referred to as a zero-violation *assignment. If one or more of the inequalities are violated then f is an assignment with constraint violations. The elements of the set F can be referred to as* frequencies *(or as* channels*).*

A zero violation assignment for the constraint graph in Fig. 1 is shown in Fig. 3. It is implicit in the above definition that the frequencies are equally spaced, although not all of the frequencies need be available. Indeed, the set of available frequencies may vary from transmitter to transmitter.

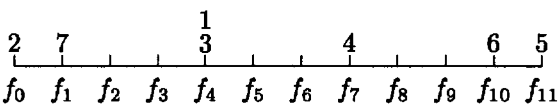


Figure 3: An Assignment for the Constraint Graph in Figure 1

Definition 3 *If all frequencies in F are available to all transmitters, and K is a minimum over all zero-violation assignments then the assignment is a minimal assignment. This minimal value of K is the* minimum span *of G, denoted $spn(G)$.*

Thus the span of an assignment is the difference between the largest channel used and the smallest channel used and $spn(G)$ is the minimum span over all possible assignments. Although the problem of minimising the span may not always be the problem that has to be solved in a particular application, it does give an idea of the spectral requirement for a network. It is also often possible to make accurate or fairly accurate assessments of the span.

The *fixed spectrum* problem starts with a set of (not necessarily consecutive) frequencies, or a number of sets of frequencies with a set defined for each transmitter. There may also be additional constraints such as fixed frequencies for certain transmitters. It is necessary to find an assignment (possibly with constraint violations) in such a way that some cost function C measuring interference

[3]Information on the CALMA project can be found on the World Wide Web at http://www.win.tue.nl/math/bs/comb_opt/hurkens/calma.html

is minimised. This cost function C can take many forms:

1. The number of constraint violations must be minimised. This cost function is often used, but has the disadvantage that there is no control over which constraints are violated if a zero cost solution is not possible, or how much they are violated by.
2. The sum of the amounts by which the constraints are violated must be minimised. This may somewhat mitigate the problem with the previous cost function, but does not remove it altogether.
3. Some overall measure of the amount by which signal–to–interference ratios at receivers are inadequate, taken over the network as a whole, must be minimised. This type of measure does appear to have some attractions, as it is closer to the real problem. The real difficulty is that it is relatively time consuming to compute. The algorithms need to explore enormous search spaces, and easily computed cost functions are necessary.
4. Some weighted function on the constraint violations must be minimised. For example, a weight w_{ijl} can be associated with the situation where transmitter i is assigned frequency f, transmitter j is assigned frequency g and $|f - g| = l$. Then w_{ijl} is zero if l is large enough to satisfy the appropriate constraint, and otherwise represents a penalty for this situation occurring. The cost function is then the sum of all the w_{ijl} over the assignment. This form of weighted cost function has been used by mobile telephone companies for some years, and has significantly increased in popularity during the past year [6]. Perhaps the main disadvantage, apart from the extra storage required for the weights in large problems, is the need to find a rational method by which the weights can be set appropriately.

Sometimes there is a mixture of *hard* constraints (which must be satisfied) and weighted (or *soft*) constraints involved in the cost function. This may, in fact, be little different from the case where all constraints are soft but some have very high weights and are effectively hard.

The weights associated with soft constraints can be determined by reasons other than those purely concerned with the incidence of interference. They may be set higher if, for example, one of the transmitters involved in the constraint handles a particularly large volume of traffic, or if some of the communication handled is particularly critical.

3 Algorithms for solving the frequency assignment problem

Since the paper of Zoellner and Beall [1] many authors have proposed algorithms for solving the generalised graph colouring formulation. It is useful to classify these into a number of different classes. Some algorithms can only be used for minimum span problems. Others can be used for both minimum span and fixed spectrum problems. Note that an algorithm for fixed spectrum problems can always be used for minimum span problems by repeatedly reducing the cost function to zero and removing the highest frequency.

In the classification which follows the algorithms will be introduced for whichever of the two problems they most directly address. In most cases, they can be modified to deal with the other problem.

3.1 Exact algorithms

Exact algorithms can be used for the minimum span problem. A number of techniques such as backtracking and forward checking [4, 7] can be used. If exact algorithms terminate, then the best possible solution is obtained. However, the frequency assignment problem is NP hard and so it can be assumed that exact algorithms will only terminate for the very smallest problems. This might typically be taken as problems with less than thirty transmitters, but it does depend on the number of frequencies available. Consequently, exact algorithms are not used in practice.

3.2 Partial backtracking

If a backtracking algorithm for a minimum span problem is to be made to work in a realistic time, the amount of backtracking must be limited. This can be done successfully, see for example [8]. However, it is unlikely that such an algorithm will have as much flexibility in exploring the search space as the best heuristic algorithms. Partial backtracking algorithms will not be considered further here.

3.3 Sequential algorithms

Sequential (or *greedy*) algorithms simply assign frequencies to transmitters with no reassignment possible. They are generally described in terms of minimum span problems. Sometimes they give reasonably good assignments. On some problems however, they are capable of very significant improvement. They are best used to give a starting assignment that can then be improved using a heuristic algorithm (see 3.9).

Sequential algorithms can make use of a number of different orderings of transmitters and of frequencies introduced by Hale [9]. Firstly the transmitters are listed in some order. Initially the first transmitter in this ordering is assigned the first frequency. Several different orderings for the transmitters were given by Hale, together with several different methods of selecting the next frequency.

3.3.1 Initial ordering of transmitters.

The initial ordering of the transmitters is an ordering of the vertices of the constraint graph corresponding to the transmitters. There are several possibilities:

Largest degree first. The transmitter are listed in decreasing order of their degree (number of incident edges) in the constraint graph.

Largest degree first (excl.). This is the same as the previous ordering except that the calculation of the degree excludes transmitters that have already been ordered, and therefore have been removed from the graph (with all their incident edges).

Smallest degree last. The transmitters of smallest degree are removed (with their incident edges). When all transmitters have been removed the list is reversed to give the final ordering.

Additional orderings. Additional orderings of the above types can be based on the generalised degree of the vertex instead of the degree. Here the number of incident edges of the vertex is replaced by the sum of the numbers ϕ_{ij} on the edges incident with the vertex. If the number of frequencies available at each vertex is not constant, there are also orderings based on the number of available frequencies.

3.3.2 Selecting the next transmitter.

The next transmitter to be assigned need not be the next unassigned transmitter in the initial ordering.

Sequential. Select the next unassigned transmitter in the initial ordering.

Generalised saturation degree. Let V be a set of transmitters and V_c be the transmitters of V already assigned frequencies. Frequency n is said to be denied to the unassigned transmitter v if there is a transmitter u in V_c assigned to frequency n such that transmitter v and u would interfere (the constraint between the frequencies given by the constraint graph is not satisfied). If frequency n is denied to transmitter v, the ***influence*** of frequency n, denoted by I_{nv}, is the largest constraint on any edge connecting v to a transmitter assigned to frequency n. The number

$$\sum I_{nv}$$

(where the sum is taken over all frequencies n denied to v) is called the ***generalised saturation degree*** of v. The technique for selecting the next transmitter is as follows: Select a transmitter with maximal generalised saturation degree (break ties by selecting the transmitter occurring first in the initial ordering).

3.3.3 Selecting a frequency.

There are four possible ways of selecting the next frequency:

Smallest acceptable frequency. Assign to the selected transmitter v the smallest acceptable frequency i.e. the lowest numbered frequency to which v can be assigned without violating any constraints.

Acceptable occupied frequency. The selected transmitter is assigned any acceptable occupied frequency (no ordering of the occupied frequencies occurs). If there is no acceptable occupied frequency, assign the transmitter to the smallest acceptable frequency.

Smallest acceptable occupied frequency. This technique attempts to minimise the number of frequencies used in the assignment. The selected transmitter is assigned to the smallest acceptable occupied frequency, if there is no acceptable occupied frequency, assign the transmitter to the smallest acceptable frequency.

Smallest acceptable most heavily occupied. The selected transmitter is assigned to the smallest acceptable most heavily occupied frequency. If there is no acceptable occupied frequency assign the transmitter to the smallest acceptable frequency.

3.3.4 Combinations of orderings

With all combinations of orderings there are 48 different sequential algorithms (or 64 orderings using the additional orderings, if not all frequencies are available at all transmitters). The best method to use is not easily predictable in advance. As the methods are all fast (except possibly for the methods involving the generalised saturation degree ordering), it seems sensible to apply several or all of these different algorithms and take the best result. Further details on sequential algorithms can be found in [9, 4]. The use of sequential algorithms in hybrids with meta-heuristics will be discussed in 3.9.

3.4 Neural networks

Several authors have demonstrated the feasibility of using neural networks to find good frequency assignments, see for example [10, 11, 12]. However, the results do not appear to be competitive with those obtained using the best meta-heuristic algorithms, so they will not be considered further here.

3.5 Genetic algorithms

Genetic algorithms have been used by several authors to find good frequency assignments. These algorithms mimic some of the processes of evolution and natural selection. More information on using them for frequency assignment can be found in [4]. Until recently the results were not competitive with the best current meta-heuristic algorithms. This was possibly due to the difficulty in defining the crossover operation in a meaningful way for frequency assignment problems. Indeed, the best approaches made use of a heuristic crossover operation based on simulated annealing or tabu search.

Very recently a new approach to the use of genetic algorithms in solving minimum span frequency assignment problems has been proposed [13]. The genetic algorithm is used to determine the best initial ordering of transmitters for use in a sequential algorithm. The results are very competitive with the current best algorithms for minimum span problems.

3.6 Hill climbing algorithms

Simple hill climbing algorithms can be used to find moderately good solutions to fixed spectrum frequency assignment problems. Essentially, assignments are replaced by neighbouring assignments if the cost function is reduced. As there is no mechanism for escaping from local minima, they should be applied many times with different starting solutions. More information on these algorithms can be found in [4].

3.7 Simulated annealing

Simulated annealing (SA) is a meta-heuristic algorithm derived from statistical mechanics. It can be used to find near minimum cost solutions in a search space with many degrees of freedom and subject to a large number of constraints. The method allows the search to proceed with the cost function reducing most of the time, but it is allowed to increase sometimes to permit escape from local minima which are not global minima.

The method itself has a direct analogy with thermodynamics, specifically with the way that liquids freeze and crystallise, or metals cool and anneal. At high temperatures, the molecules of a liquid move freely with respect to one another. If the liquid is cooled slowly, thermal mobility is restricted. The atoms are often able to line themselves up and form a pure crystal that is completely regular. This crystal is the state of minimum energy for the system, which would correspond to the optimal solution in a mathematical optimisation problem. However, if a liquid metal is cooled quickly i.e. quenched, it does not reach a minimum energy state but a somewhat higher energy state corresponding, in the mathematical sense, to a suboptimal solution found by iterative improvement or hill-climbing.

In order to make use of this analogy, the following elements must be present:

1. A description of the *configuration* $\mathbf{X} = (x_1, x_2, ..., x_N)$ that represents a solution; for the frequency assignment problem this is an assignment of frequencies to each of N transmitters.

2. A generator of random changes to the configuration. These changes are typically solutions in the *neighbourhood* of the current configuration, for example, a change in one of the parameters, x_i. Thus in this case the frequency assigned to one transmitter might change.

3. A cost function $C(\mathbf{X})$ whose minimisation is the goal of the simulated annealing procedure. This is the analogue of the energy in the thermodynamic analogy.

4. A control parameter t and an *annealing schedule* which indicates how t is lowered from high values to low values. This is the analogue of temperature in the thermodynamic analogy.

Suppose that as a result of a random change to the configuration, the cost function changes from C_{old} to C_{new}. If $C_{new} < C_{old}$ the new configuration is always accepted. If $C_{new} > C_{old}$, then the new configuration may still be accepted, with probability $e^{(-(C_{new}-C_{old})/Bt)}$, where B is a fixed constant known as the Boltzmann constant. This general scheme, of always taking a downhill step while sometimes taking a uphill step is known as the Metropolis Algorithm. The simulated annealing procedure of Kirkpatrick et al. [14] uses the Metropolis Algorithm but *varies* the temperature parameter t from a high value at which most new configurations are accepted to a low value t_{min} at which no new configurations are accepted. Let “random” denote a random number in the range

[0,1] and let NUM_{loop} be the number of random changes in configuration at each temperature t. Also, let C denote one of the cost functions C for fixed spectrum problems described in section 2. Then the full simulated annealing procedure for minimisation can be written as follows:

Initialise t

Generate random configuration $\mathbf{X_{old}}$

WHILE $t > t_{min}$ DO

FOR $i = 1$ to NUM_{loop} DO

generate new configuration, $\mathbf{X_{new}}$, from $\mathbf{X_{old}}$

calculate new energy, C_{new}

calculate $\Delta C = C_{new} - C_{old}$

IF $\Delta C < 0$ or *random* < *prob* = $e^{-\Delta C/t}$ THEN

$\mathbf{X_{old}} \Leftarrow \mathbf{X_{new}}$

$C_{old} \Leftarrow C_{new}$

END IF

END FOR

reduce t (e.g. $t = 0.9t$)

END WHILE

It will now be described in more detail how an assignment is represented, how new configurations are generated and how the temperature reduces, the so called *cooling schedule*

3.7.1 Representation of an Assignment

Suppose that the available frequencies are denoted $(d_1, d_2, d_3, \ldots)$. Let $f_j = d_{x_j}$. Then the frequency assignment $(f_1, \ldots, f_N)$ can be represented by an array of indices $[x_1, \ldots, x_N]$.

3.7.2 Generation of New Configurations

Three different generators can be used to produce a new assignment from the current assignment.

Single Move. Here the neighbours of an assignment are those assignments where the array of indices differs in precisely *one* component. Thus if the new assignment is represented by $[x'_1, \ldots, x'_N]$ then $(f'_1, \ldots, f'_N)$ is a neighbour of $(f_1, \ldots, f_N)$ if there exists j, $1 \leq j \leq N$, such that $x'_j \neq x_j$, and $x'_i = x_i$ for all $i = 1, \ldots, N$ with $i \neq j$. If there are F frequencies available, then any assignment has $N(F-1)$ neighbours.

Double Move. Here the neighbours of an assignment are those assignments where the array of indices differs in precisely *two* components.

Restricted Single Move. For a given assignment, transmitters which are assigned frequencies which violate one or more of the constraints are referred to as *violating transmitters*. The *restricted single move* generator involves randomly selecting a violating transmitter i and setting $x_i = x'_i$ where $1 \leq x'_i \leq F$. This move generator has been found to be particularly effective when the number of violations is small. Thus it should always be used when the hybrid of sequential algorithms and simulated annealing (to be described in 3.9) is used. It may well be a good choice for all applications.

3.7.3 Starting and finishing temperatures

The starting temperature can be determined by first setting $t_0 = 1$ and performing 100 iterations of the FOR loop from the main loop of the algorithm in 3.7. If the acceptance ratio, χ, defined as the number of accepted trial assignments divided by 100 (NUM_{loop}), is less than 0.9, double the current value of t_0. Continue this procedure until the observed acceptance ratio exceeds 0.9 (with χ reinitialised to zero prior to starting the FOR loop).

The algorithm terminates when the temperature, t_k, falls below a user specified value t_{min}, or the number of *frozen temperatures* exceeds a chosen value, usually 10. A frozen temperature occurs when no new assignments are accepted for a given temperature t_k (i.e. after NUM_{loop} iterations of the FOR loop).

Annealing Schedule (reduction of t_k)

Three different annealing schedules will be described. In all three cases the parameter NUM_{loop} is set to N, the number of transmitters.

Cooling 1 (geometric):

With a simple cooling scheme t_k is reduced according to the formula:

$$t_{k+1} = \alpha t_k, \qquad \text{with } \alpha \in [0,1]$$

Cooling 2 (Costa [15])

The geometric scheme (cooling 1) decreases the temperature by the specification of the parameter α; the number of iterations at each temperature is fixed at N. In the cooling 2 scheme, the temperature reduction is again specified by the parameter α, however the number of iterations at each temperature is specified by

$$N_{k+1} = \lceil N_k/\alpha \rceil$$

(up to a maximum of $2000N$), where $N_0 = N$, the number of transmitters.

As the temperature decreases the number of trial assignments tested at each temperature increases. This scheme was found to be particularly effective in [15].

Cooling 3 (Hurley and Smith [16]):

A more complex scheme which is slower, but gives better results, is when the parameter t_k is calculated using:

$$t_{k+1} = t_k \cdot \left(1 + \frac{\ln(1+\delta)\cdot t_k}{3\sigma(t_k)}\right)^{-1}$$

where δ is set to 0.1 and

$$\sigma(t_k) = \frac{1}{NUM_{loop}} \cdot \sqrt{\Phi}$$

where

$$\Phi = NUM_{loop} \cdot \sum_{i=1}^{NUM_{loop}} \left({C_i}^k\right)^2 - \left(\sum_{i=1}^{NUM_{loop}} {C_i}^k\right)^2 + 0.5$$

and where ${C_i}^k$ is the cost function value for the assignment obtained at iteration i, at temperature t_k.

3.8 Tabu search

The basic idea of the tabu search meta-heuristic (TS) [17, 18] is to explore the *search space* of all feasible solutions by a sequence of *moves*. A move from one solution to another is generally the best available. However, in order to prevent cycling and to provide a mechanism for escaping from locally optimal but not globally optimal solutions, some moves, at one particular iteration, are classified as forbidden or *tabu* (or *taboo*). Moves are regarded as tabu by consideration of the short-term and long-term history of the sequence of moves. A very simple use of this idea might be to classify a move as tabu if the reverse move has been made recently or frequently. There are also ***aspiration criteria*** which override the tabu moves in particular circumstances. These circumstances might include the case when, by forgetting that a move is tabu, a solution which is the best so far is obtained.

Suppose that it is required to minimise some cost function C on a search space S. For combinatorially hard problems, such as fixed spectrum frequency assignment problems, it may only be possible to obtain sub-optimal solutions, in which C is close to its minimum value. Sub-optimal problems may be obtained when a certain threshold for an acceptable solution has been achieved or when a certain number of iterations have been completed.

A characterisation of the search space S for which tabu search can be applied is that there is a set of k moves $M = \{m_1, ..., m_k\}$ and the application of the moves to a feasible solution $s \in S$ leads to k, usually distinct, solutions $M(s) = \{m_1(s), ..., m_k(s)\}$. The subset $N_{set}(s) \subseteq M(s)$ of ***feasible*** solutions is known as the ***neighbourhood*** of s.

The method starts with a (possibly random) solution $s_0 \in S$ and determines a sequence of solutions $s_0, s_1, ..., s_n \in S$. At each iteration, $s_{j+1} (0 \leq j < n)$ is selected from the neighbourhood $N_{set}(s_j)$. The selection process is first to determine the tabu set $T_{set}(s_j) \subseteq N_{set}(s_j)$ and the aspirant set $A_{set}(s_j) \subseteq T_{set}(s_j)$. Then s_{j+1} is the neighbour of s_j which is either an aspirant or not tabu and for which $C(s_{j+1})$ is minimal; that is, $C(s_{j+1}) \leq C(s')$ for all $s' \in (N_{set}(s_j) - T_{set}(s_j)) \cup A_{set}(s_j)$.

To apply tabu search to fixed spectrum frequency assignment problems with a cost function C as described earlier, it is necessary to decide how a solution is represented, how the neighbourhood of a solution is determined, how moves are classified as tabu and the aspiration criteria.

Representation of an Assignment. This is the same as already described for simulated annealing.

Neighbourhood Structure: Full Neighbourhood. The full neighbourhood includes all the neighbours of an assignment f produced by using the ***single move*** generator described for the simulated annealing algorithm.

Neighbourhood Structure: Restricted Neighbourhood. The reduced neighbourhood consists of all neighbours of an assignment f produced by using the ***restricted single move*** generator described for the simulated annealing algorithm. This increases the efficiency of the algorithm since this restricted neighbourhood will tend to be much smaller than the full neighbourhood and the algorithm concentrates on changes which affect constraint violations. However, if the restricted neighbourhood consists of moves which are tabu and do not satisfy the aspiration criteria, the full neighbourhood will be selected for this iteration. Use of the restricted neighbourhood has been found to be effective for frequency assignment, particularly when C is small. Thus it should always be used when the hybrid of sequential algorithms and tabu search for minimum span problems (to be described in 3.9) is used. It may well be a good choice for all applications.

Restricted Random Neighbourhood. Here the neighbourhood of an assignment f is generated by randomly selecting a violating transmitter and randomly assigning a different frequency.

Definition of a Tabu Move: Short and Long Term Memory. A move to a neighbour (i, x_i') corresponding to changing the assignment of transmitter i to x_i' is said to be *tabu* if it does not satisfy short term or long term memory conditions. These conditions are determined by two numbers: a positive integer r and a real number β, with $0 < \beta < 1$. The short term memory condition specifies that any transmitter, i, cannot be assigned the same frequency over any of the previous r moves. The long term memory condition specifies that the proportion of the number of times transmitter i had been changed over all iterations does not exceed β. Thus, if at iteration j, the frequency $d_{x_i j}$ assigned to transmitter i has changed, then a move at iteration $k+1$ is tabu if either (short term memory)

$$d_{x_i k+1} = d_{x_i t} \qquad \text{where } k - r < t \leq k$$

or (long term memory)

$$\frac{1}{k} \sum_{j=1, i_j = i_{k+1}}^{k} 1 \quad > \quad \beta$$

At each iteration, the method selects from the non-tabu neighbours that neighbour $f^{(k+1)}$ of $f^{(k)}$ for which $C(f^{(k+1)})$ is minimal. Note that it is possible that $C(f^{(k+1)}) > C(f^{(k)})$; this allows escape from local minima.

Tabu search allows a tabu move to be selected when certain aspiration criteria are satisfied. A tabu move can be selected if the neighbour $f^{(k+1)}$ satisfies

$$C(f^{(k+1)}) \leq C(f')$$

for all neighbours f' of $f^{(k)}$ and

$$C(f^{(k+1)}) < C(f^{(j)}) \qquad \forall\, j, \quad 1 \leq j \leq k.$$

In summary, if f_{tabu} denotes the best tabu neighbour of $f^{(k)}$ and f_{non} the best non-tabu neighbour, then the rule for determining $f^{(k+1)}$ is:

if $C(f_{tabu}) < C(f_{non})$ and
$C(f_{tabu}) < C(f^{(j)})$ for $1 \leq j \leq k$
then $f^{(k+1)} = f_{tabu}$
else $f^{(k+1)} = f_{non}$

Notice that there are some implicit relations between the parameters r and β which have to be satisfied. If $r > N$ or $\beta < 1/N$, then, after N non-tabu moves, every move is tabu. Also, after k iterations of non-tabu moves, it follows from the recency condition that a transmitter can have changed at most $\lceil k/r \rceil$ times. Hence, the long term memory value, to have any effect, should satisfy

$$k\beta < \lceil k/r \rceil$$

Therefore a relation between r and β is obtained, which is

$$1/N \leq \beta \leq 1/r$$

In practice, a good choice is

$$\beta = \lambda/r + (1 - \lambda)/N$$

for some value of λ with $0 \leq \lambda \leq 1$.

Suitable values for the short term memory and long term memory conditions are $r = \frac{2*N}{5}$ and $\lambda = 0.5$ respectively.

The algorithm is terminated if the cost function C becomes zero or if a pre-determined maximum number of iterations is reached.

3.9 The use of SA and TS in Minimum Span Frequency Assignment

It is possible to use either the SA or TS algorithm for minimum span frequency assignment. A random starting assignment can be used with a set of consecutive frequencies for which the span is larger than some estimate of the minimum span. The appropriate meta-heuristic algorithm is then used to reduce the cost function to zero. The largest frequency is removed and transmitters assigned this frequency are reassigned randomly (or in some optimum way as in [19]). The cost function is then reduced to zero again and the process repeated. When it is no longer possible to reduce the cost function to zero, the previous zero cost assignment gives the best span obtained.

It turns out that it is much better to use an initial assignment obtained from a sequential algorithm (or preferably the best of many sequential algorithms) instead of a random assignment. The idea was first suggested by Costa [15]. Fuller details can be found in [4].

3.10 Using subgraphs in frequency assignment

It has been demonstrated in [20] (see also [21]) that lower span assignments can sometimes be obtained by identifying some subgraph of the constraint graph. The subgraph is often a clique (a maximal complete subgraph) of the constraint graph, or a clique with some additional vertices. The subgraph is assigned first, using one of the meta-heuristic algorithms described previously. The assignment is fixed and then extended to an assignment of the full constraint graph, using one of the meta-heuristic algorithms. Sometimes it is then beneficial to unfix the subgraph and continue to try to improve it using one of the meta-heuristic algorithms. The appropriate subgraph is often a subgraph that is used to obtain the best lower bound for the span, and will be dealt with in more detail in section 4. The method does not always prove effective, but when it is effective significantly better assignments are obtained.

If a fixed spectrum problem has to be solved where the available spectrum only lacks a few frequencies from those used in a minimum span assignment, it may be helpful to work from the minimum span assignment, rather than apply the fixed spectrum method directly. Transmitters assigned unavailable frequencies are reassigned randomly (or in some optimum way as in [19]). The meta-heuristic is then applied to this assignment to reduce the cost further. Sometimes better assignments are obtained this way than can be obtained directly, particularly if the initial minimum span assignment was a particularly good assignment obtained using subgraphs. However, the approach seems hardly likely to be effective if the number of available frequencies is significantly less than the number used in a minimum span assignment.

4 Evaluation of algorithms and assignments

Definition 4 *An assignment which is a solution of a minimum span problem will be called an* optimal assignment *if its span is equal to one of the lower bounds which will be described in 4.1, 4.2 and 4.4. An assignment which is a solution of a fixed spectrum problem will be called optimal if its cost function C is equal to some lower bound.*

Questions about the effectiveness of frequency assignment methods and algorithms fall logically into two classes. The first type of question assumes a given problem is defined using binary constraints, by giving a constraint graph as in section 2, for example. It is then asked how close to optimal a particular assignment is with respect to this problem definition. More generally, it can be asked how good a particular algorithm is at generating optimal or near optimal solutions given such problem definitions. The second type of question asks whether a definition in terms of binary constraints is really the best way to represent the problems presented by the real world. Although representations involving only binary constraints (or possibly binary constraints and intermodulation products) are universally popular, they may not be the best way of achieving the aim of minimising interference while conserving spectrum. The second type of question will be addressed in section 5.

Many papers have presented an algorithm for the frequency assignment problem and tested the algorithm on datasets which are not available to others. This can make the effectiveness of the algorithm very difficult to judge. There are two principal datasets available which can be used to test the relative merits of different algorithms. The first is the CELAR dataset already mentioned in section 2 in connection with the CALMA project. Unfortunately some of the constraints are not addressed by the majority of published algorithms. The second dataset is made up of the "Philadelphia" problems. In [4, 20] it was shown that the simulated annealing and tabu search meta-heuristics in the package FASOFT, when using the techniques of 3.9 and 3.10 and the second form of the cost function C defined in section 2, are capable of solving the main variations of the Philadelphia problems to optimality. The Philadelphia problems are cellular problems. They may not be altogether typical of the problems that arise in reality in modern systems. Additionally, some researchers have inadvertently tested their algorithms on rather easy variations of the Philadelphia problems. For example, if the cosite value is increased from 5 to 7 the problems become easy for any reasonable algorithm.

Thus there is a need for further standard datasets of a demanding nature to be made available to researchers via the world wide web. The same algorithms in FASOFT have been shown to be effective on datasets supplied by several civil and military network operators. Generally, for problems derived from cellular telephone networks, the algorithms give solutions for minimum span problems ranging from optimality to within 5% of optimality. Optimal solutions have been found for problems with up to 10,000 transmitters using FASOFT [4]. For military radio link problems solutions range from optimality to within 10% of optimality. Of course, it is not always easy to obtain example datasets from operators in order to generate such estimates. Here the civil estimates are based on data from three operators. These results can be considered very satisfactory given the known difficulty of the frequency assignment problem.

The classification of solutions as optimal or near optimal can only be done if good lower bounds are available. The generation of good lower bounds for minimum span problems [22, 23, 24, 25] will be considered first. Then the possibility of generating such bounds for the cost function in fixed spectrum problems will be mentioned.

4.1 Clique bounds

The most common lower bound for the span in frequency assignment problems is based on cliques. The idea is borrowed from the theory of colourings of graphs. A *clique* of a graph G is a maximal complete subgraph of G. Thus every pair of vertices of the subgraph are adjacent, and the subgraph is not contained in any larger such subgraph (some authors omit this maximality requirement). A clique can be regarded as a set of transmitters for which there is a constraint between the frequencies assigned to any pair. In many applications cliques tend to correspond to clusters of geographically close transmitters. It is possible to define several different levels of clique, corresponding to different minimum edge labels:

Definition 5 *A* level-p clique *of G is a complete subgraph in which every edge has label at least p, and which is not contained in any larger such complete subgraph.*

Thus a level-1 clique, for example, corresponds to a set of transmitters for which every pair of transmitters cannot be assigned the same channel, or a first adjacent channel.

Theorem 1 *If C_p is a level-p clique of a constraint graph G then*

$$spn(G) \geq (p+1)(|V(C_p)|-1).$$

where $V(C_p)$ denotes the vertex set of the clique C_p.

Proof The minimum span of G cannot be less than the minimum span of the subgraph C_p of G. For any chosen minimal span assignment f of C_p renumber the vertices of C_p as $v_0, v_1, \ldots, v_{|V(C_p)|-1}$ in ascending order of the channel assigned to them. The span of the assignment is the difference between the largest and the smallest channel used, i.e.

$$\begin{aligned} spn(C_p) &= f(v_{|V(C_p)|-1}) - f(v_0) \\ &= \sum_{j=0}^{|V(C_p)|-2} f(v_{j+1}) - f(v_j) \\ &\quad \text{(as all but two of the terms cancel in pairs)} \\ &\geq \sum_{j=0}^{|V(C_p)|-2} p+1 \\ &= (p+1)(|V(C_p)|-1). \end{aligned}$$

□

Example 1 *The constraint graph shown in Fig. 1 has minimum span 11. A minimum span assignment is shown in Fig. 3. Theorem 1 applied to the level-3 clique {2 3 5} gives $spn(G) \geq 8$. Similarly, applying the theorem to the level-2 clique {2 3 4 5} gives $spn(G) \geq 9$. It will be seen later that the clique bound is capable of improvement for this example.*

4.2 Travelling Salesman Bounds

The clique bound can sometimes be improved by making use of *Hamiltonian paths* [23, 24]. A Hamiltonian path in a graph G is a path through all of the vertices of the graph which visits each vertex once and once only. The description of the bound is more convenient if a new graph is constructed from G. G' is a weighted complete graph which has the same vertices as G and has the weight c_{ij} of each edge v_iv_j of G' given by:

$$\begin{aligned} c_{ij} &= 0 && \text{if } v_iv_j \text{ is not an edge of } G, \\ c_{ij} &= \phi_{ij}+1 && \text{if edge } v_iv_j \text{ has label } \phi_{ij} \\ & && \text{in } G\ (\phi_{ij}=0,1,\ldots). \end{aligned}$$

$H(G')$ denotes the total weight of a minimum weight Hamiltonian path in G'.

Theorem 2 *For a given constraint graph G:*

$$spn(G) \geq H(G')$$

Proof For any chosen minimal span assignment f of G number the vertices of G as $v_0, v_1, \ldots, v_{|V(G)|-1}$ in ascending order of the channel assigned to them (and arbitrary order for vertices assigned the same channel). Then $v_0, v_1, \ldots, v_{|V(G)|-1}$ is a Hamiltonian path H_f in G'. The span of the assignment is the difference between the largest and the smallest channel used, i.e.

$$\begin{aligned} spn(G) &= f(v_{|V(G)|-1}) - f(v_0) \\ &= \sum_{j=0}^{|V(G)|-2} f(v_{j+1}) - f(v_j) \\ &\quad \text{(as all but two of the terms cancel in pairs)} \\ &\geq \sum_{j=0}^{|V(G)|-2} c_{v_{j+1}v_j} \\ &= \text{the total weight of } H_f \\ &\geq H(G'). \end{aligned}$$

□

The problem of determining the minimum weight of a Hamiltonian path in a graph is usually known as the open, symmetric travelling salesman problem [26]. Thus the bound of Theorem 2 can be referred to as the *travelling salesman bound.*

The bound of Theorem 2 should be applied to a subgraph of the constraint graph and not to the constraint graph itself. In most problems the application of the theorem to the full constraint graph gives a lower bound which is too small to be useful. In order to obtain a strong bound, a suitable choice of subgraph is a clique or a clique with some vertices added. The question of how this subgraph should be obtained will be considered in 4.3.

A method due to Volgenant and Jonker [26] and software described by Volgenant [27] [1] can be used to calculate the travelling salesman bound. The software generally gives satisfactory results when the subgraph has up to 250 vertices. Even when the algorithm does not converge in reasonable time, the difference between the lower and upper bound is small, so a good lower bound for $H(G')$ is usually obtained.

Another bound, which is easier to calculate, is the *spanning tree bound.* A spanning tree in a graph is a connected subgraph of the graph which contains every vertex and no cycles. Let $S(G')$ denote the total weight of a minimum weight spanning tree in G'. As a Hamiltonian path is a spanning tree, it follows that $H(G') \geq S(G')$ and so the following result is immediate:

Theorem 3 *If* G *is a constraint graph then*

$$spn(G) \geq S(G')$$

As with the travelling salesman bound, theorem 3 should normally be applied to a suitable subgraph and not to the full constraint graph. The spanning tree bound may not be as strong as the travelling salesman bound, but is much easier to calculate. When applied to a clique it can be stronger than the clique bound. A simple greedy algorithm, known as Prim's Algorithm [28], can be used to find $S(G')$ with no restriction on the size of subgraph.

Returning to Example 1 it can be seen that applying both the spanning tree bound and the travelling salesman bound to the clique {2 3 4 5} give lower bounds of 11, which equals $spn(G)$.

4.3 Finding cliques

It has already been pointed out that the use of the travelling salesman and spanning tree bounds only gives good results it they are applied to a suitably chosen subgraph, and not to the full constraint graph. The simplest way to do this is to apply one of the bounds to the largest level-p clique for all appropriate values of p, and then choose the value of p that gives the best bound.

An algorithm is necessary to find maximal cliques. The exact algorithm of Carraghan and Pardalos [29] gives good results for frequency assignment problems up to 700–800 transmitters. It can, however, be important that a good ordering of the transmitters is used for larger problems. The orderings discussed in 3.3 are suitable. For problems with more than about 800 transmitters a heuristic maximal clique algorithm can be used.

Sometimes it is possible to add a small number of vertices to a clique and the travelling salesman bound increases initially. One way to do this is to assign the clique and determine the number of available frequencies for each transmitter not in the clique. The transmitter with the smallest number of available frequencies is added to the clique and then the process is repeated with the subgraph obtained. A fuller discussion of this topic can be found in [20].

4.4 Mathematical programming

Mathematical Programming techniques can sometimes be used to improve the bounds described in 4.2. An integer programming formulation of the travelling salesman problem can be formulated as follows. Form the graph G'_0 from G' by adding a dummy vertex v_0 joined by an edge of weight 0 to every vertex of G'. This dummy vertex v_0 converts the open symmetric travelling salesman problem to a closed symmetric travelling salesman problem, in which the requirement is to minimize the length of a circuit instead of the length of a path. If a minimal weight circuit is found in G'_0 then removal of the vertex of weight 0 gives a minimum weight path in G'. $H(G')$ is then equal to the solution of the following well known integer program for the closed symmetric travelling salesman problem (TSP):

$$\text{Minimize} \sum_{v_i v_j \in E(G'_0)} c_{ij} x_{ij} \quad (1)$$

$$\text{subject to} \sum_{j: v_i v_j \in E(G'_0)} x_{ij} = 2; \; v_i \in V(G'_0) \quad (2)$$

$$\sum_{v_i \in S, v_j \in V(G'_0) \setminus S} x_{ij} \geq 2; \; S \subset V(G'_0) \quad (3)$$

$$x_{ij} \in \{0,1\}; \qquad v_i v_j \in E(G'_0). \quad (4)$$

The formulation gives a minimum weight Hamiltonian circuit in G'_0 from which a minimal

[1] The software is available on the World Wide Web. See: http://www.mathematik.uni-kl.de/~ wwwwi/WWWWI/ORSEP/contents.html
ftp://www.mathematik.uni-kl.de/pub/Math/ORSEP/VOLGENAN.ZIP

weight Hamiltonian path of weight $H(G')$ can be obtained by deleting the dummy vertex v_0. The integer variable x_{ij} is equal to 1 if edge v_iv_j is in the Hamiltonian circuit and is 0 otherwise. The total weight of the circuit to be minimised is given as 1. The equations 2 represent the requirement that there are two edges of the Hamiltonian circuit at each vertex. The inequalities 3 are called the subtour elimination constraints. These ensure that a single circuit is obtained, rather than the union of several disjoint circuits.

If this integer program can be solved exactly the value of $H(G')$ is obtained. However, this may not be practical. One particular difficulty here is the memory required to store the subtour elimination inequalities. An approach to making a solution practical is to relax the problem by weakening or removing one or more of the constraints. If the constraints are weakened it may be possible to find a solution of smaller total weight. Then $H(G')$ is no longer found, but instead a lower bound for $H(G')$ is obtained, which may be still be adequate for the purpose of deriving a strong lower bound for $spn(G)$.

Relaxation of the integrality constraint 4 to

$$0 \leq x_{ij} \leq 1 \qquad v_iv_j \in E(G_0') \qquad (5)$$

gives the linear (LP) relaxation of the integer program. This relaxation can be easy to solve and is generally a good lower bound. For random graphs the bound is, on average, within 1% of the exact value of $H(G')$. However, the memory requirement problem for the subtour elimination constraints remains.

Alternatively, the integer program can be relaxed by removing the subtour elimination constraints 3. An integer program for the minimum weight perfect two-matching problem (PTMP) in G_0' is obtained. A perfect-two matching is a union of one or more circuits containing every vertex once and once only. A lower bound can also be obtained by replacing the integrality constraint 4 by the constraint 5 and using a Linear Programming package. For many subgraphs of constraint graphs consisting of a clique or a clique with some additional vertices, this lower bound is close to $H(G')$.

This LP relaxation of the minimum weight perfect two-matching problem, applied to a suitable subgraph, can be a simple, fast and robust method of obtaining good lower bounds for $spn(G)$ in circumstances where the travelling salesman bound is strong. It is never stronger than the travelling salesman bound without the addition of further constraints.

There are two circumstances in which improvements to the linear programming approach described above are possible. The first circumstance is when the travelling salesman bound is inherently weak. This can arise because the bound takes no account of constraints between non-consecutive vertices in the Hamiltonian path. The second circumstance is where there is some uncertainty over what is the best subgraph to use for the bound. The value of the travelling salesman bound (and most other bounds) reduces rapidly when more than a small number of critical vertices are added to the appropriate clique. This behaviour can be mitigated if the bound is improved, which may be possible by adding additional constraints, referred to as *frequency assignment constraints.*

Associate a non-negative integer variable e_{ij} with each edge v_iv_j of the constraint graph G. These variables e_{ij} are chosen so that there will be no constraint violations when an assignment is constructed from a Hamiltonian path $\{v_{i_1}, \ldots, v_{i_n}\}$ by setting

$$\begin{aligned} f(v_{i_1}) &= 0 \\ f(v_{i_j}) &= f(v_{i_{j-1}}) + c_{i_{j-1}i_j} + e_{i_{j-1}i_j} \\ &\quad \text{for } j = 2, \ldots, n. \end{aligned}$$

Then constraints between consecutive vertices on the Hamiltonian path no longer have to be met exactly, allowing constraints between non-consecutive vertices to be satisfied. In this case the value e_{ij} is referred to as the *excess* on the edge v_iv_j.

The formulation of the frequency assignment constraints uses the following definitions. If P is a path $v_{i_1}, v_{i_2}, \ldots, v_{i_k}$ with edge set $E(P)$, then let $X_P = x_{i_1i_2} + \cdots + x_{i_{k-1}i_k}$ and $E_P = e_{i_1i_2} + \cdots + e_{i_{k-1}i_k}$. Define the *deficit* of P as

$$d(P) = c_{i_1i_k} - (c_{i_1i_2} + \cdots + c_{i_{k-1}i_k}).$$

Let $\mathcal{P}(G')$ be the set of paths P of G' with $d(P) > 0$. Then, if $P \in \mathcal{P}(G')$ it is required that

$$X_P - (|E(P)| - 1) \leq \frac{E_P}{d(P)}.$$

If $X_P \leq (|E(P)| - 1)$ then E_P is unconstrained. If $X_P = |E(P)|$ (that is, if all edges of P are included in the Hamiltonian path), then the total excess on P must be at least as large as the deficit of P to ensure that the constraint between the end vertices of P is satisfied. This gives the following integer programming formulation of the frequency assignment problem (FAP):

$$\text{Minimize} \sum_{v_i v_j \in E(G'_0)} c_{ij} x_{ij} + \sum_{v_i v_j \in E(G')} e_{ij}$$

subject to:

$$\sum_{j: v_i v_j \in E(G'_0)} x_{ij} = 2; \quad v_i \in V(G'_0) \quad (6)$$

$$\sum_{v_i \in S, v_j \in V(G'_0) \setminus S} x_{ij} \geq 2; \quad S \subset V(G'_0) \quad (7)$$

$$X_P - (|E(P)| - 1) \leq \frac{E_P}{d(P)} \quad P \in \mathcal{P}(G') \quad (8)$$

$$x_{ij} \in \{0, 1\}; \quad v_i v_j \in E(G'_0) \quad (9)$$

$$e_{ij} \in \{0, 1, \ldots, c_{max}\}; \quad v_i v_j \in E(G') \quad (10)$$

where $c_{max} = \max_{\{i,j\}} c_{ij}$. Note that when minimizing 4.4, e_{ij} will equal zero if $x_{ij} = 0$.

A dummy vertex is used, as with the formulation of the travelling salesman problem, to express the problem in terms of a circuit rather than a path. The objective function in 4.4 is the total actual length of the path to be minimised. Equations 6 ensure that there are two edges of the circuit at each vertex. Inequalities 7 are the subtour elimination constraints. Inequalities 8 ensure that there are no constraint violations between non-consecutive vertices in the path. The fact that the variables x_{ij} are integers is expressed by 9 and the fact that the excesses are integers and at most equal to the maximum constraint value is expressed by 10.

If this integer program can be solved, it gives an exact solution of the channel assignment problem for the full constraint graph G. Although this is rarely possible, lower bounds can be derived as follows:
The integrality constraints 9, 10 are replaced by

$$0 \leq x_{ij} \leq 1; \quad v_i v_j \in E(G'_0) \quad (11)$$

$$0 \leq e_{ij} \leq c_{max}; \quad v_i v_j \in E(G') \quad (12)$$

and a linear programming relaxation of the channel assignment problem is obtained. By also omitting the subtour elimination constraints 7, a linear programming relaxation of the formulation of the perfect two matching problem (PTMP), with additional frequency assignment constraints (PTMP+FAP) is obtained. This gives solutions in an acceptable time and, when practicable, appears to give an excellent bound when applied to a suitable subgraph. If there are too many constraints in equation 8 then a subset of them must be chosen. This can be done by, for example, limiting the length of the paths $P \in \mathcal{P}(G')$ to 3 or 2.

The decision of which FAP constraints 8 to keep and which to eliminate may be critical if a strong bound is to be obtained. However, it appears that elimination of the subtour elimination constraints and relaxation of the integrality constraints makes little or no difference to the strength of the lower bound, for real frequency assignment problems. A level-0 clique is often a good candidate subgraph for this method, rather than a higher level clique. This is because the constraints in equation 8 tend to prevent constraint violations between non-consecutive vertices in the Hamiltonian path, which often occur if a travelling salesman bound is used with this level of clique.

Some results comparing this lower bound with previous methods and illustrating the improvement are given in [25]. The method can also be considerably simplified for cellular problems [24].

4.5 Lower bounds for fixed spectrum problems

The lower bounding techniques discussed so far apply only to the span in minimum span problems. It is of equal importance to derive lower bounds for the cost function in fixed spectrum problems. Without such lower bounds it is only possible to argue that fixed spectrum algorithms are effective because they can be used in the hybrid methods outlined in 3.9 to produce good solutions to minimum span problems. However, fixed spectrum problems can contain features such as illegal frequencies which are not present in the minimum span problem as defined here. It is conceivable that a fixed spectrum algorithm could work well except, for example, when there were illegal frequencies.

A small first step towards deriving lower bounds for fixed spectrum problems is taken in [24]. However, it appears that the derivation of tight bounds may be much more difficult for fixed spectrum problems than it is for minimum span problems. The reason for this is that the bounds for minimum span problems are derived "locally", from a subgraph which usually corresponds to a cluster of close transmitters. In fixed spectrum problems in which the available frequencies are significantly fewer than is indicated by the minimum span, the constraint violations may be spread over the entire network. Thus a "global" method of finding a lower bound on the cost function C is required.

4.6 Algorithms for special problems

In addition to the algorithms described in section 3, there are many other, mainly heuristic, algorithms which have been described in the literature. Indeed there are very many algorithms which have never been published. It is often very difficult to provide an assessment from published information of the performance of many of these algorithms.

The general comments about the effectiveness of algorithms in section 3 and in 4 are based on assessments against lower bounds on both civil and military data which can be considered typical of the frequency assignment problems that arise in those areas. However, many algorithms not presented in section 3 may have been developed for specific problems which may have specific characteristics. They may perform better for those problems than general purpose algorithms. Developers of such algorithms for minimum span problems are able to test their effectiveness against lower bounds, and to compare performance with general purpose algorithms which are known to work well. Ideally, the same should be true for fixed spectrum problems.

5 Evaluating the binary constraint representation

In section 2 the representation of frequency assignment problems as generalised graph colouring problems, using binary constraints, was introduced. The constraints describe the necessary channel separations to sufficiently attenuate interfering signals. The representation essentially makes a *single interferer assumption.* At each receiver the interfering signals from transmitters other than the wanted transmitter are considered individually, and an appropriate necessary separation between frequencies assigned to the wanted and the interfering transmitter is derived. This is a simplification. It would be more accurate to consider interference from all potentially interfering transmitters when deciding whether the receiver met the required minimum signal-to-interference ratio. The assumption has allowed the algorithms and theory of graph colourings to be brought to bear on the problem with considerable success. However, the meta-heuristic algorithms described in section 3 are not limited to binary constraints. Indeed, they are often used with constraints involving intermodulation products etc.. Normally only a change to the cost function is necessary. In the case of equality constraints, for example, a change to the solution representation may be necessary if the algorithm is to work effectively.

It has recently been suggested (see [30] for example) that better assignments may be obtained if the single interferer assumption is abandoned and non-binary constraints are used in the problem representation. This would of course, present a number of difficulties. It is not clear how the constraints should be derived, and the memory required for all possible non-binary constraints could be prohibitive. Thus a selective approach may be necessary.

The usual method of generating binary constraints will be considered first. Then the typical form of some non-binary constraints will be outlined. Finally, some experiments using a "constraint free" approach will be described. These experiments aim to assess the relative merits of the binary and non-binary methods.

The derivation of binary constraints might typically proceed as follows. For each transmitter T_i, either there is a single receiver or some typical receiver positions are identified. For each such receiver the wanted signal strength is predicted using some propagation model. For each potentially interfering transmitter T_j, the interfering signal at the same receiver is predicted, using the same propagation model. The channel separation necessary for this receiver to have an adequate signal-to-interference ratio is determined. The process is then repeated for a single receiver or some typical receiver positions of T_j with T_i as the interfering transmitter. The necessary separation between f_i and f_j is then the largest of these separations, taken over all receivers of T_i and all receivers of T_j. On the assumption that there is only a single dominant interferer at each receiver location, the signal-to-interference ratio at that location is taken to be adequate.

Suppose now that instead of considering a single source of interference at each receiver, several of the strongest potential interferers are considered. If the interference is considered to be additive, there may be some choices for the possible channel separations. A smaller separation might be allowable for one interferer if a larger separation is imposed for another. The form that the constraints might take is illustrated in Example 2.

Example 2 *Suppose the $R_{7\ 3}$ denotes the third receiver position of transmitter 7. Then consideration of the interference from the worst potential interferers $T_2, T_6, T_9, T_{11}, T_{12}$ leads to the following table of possible separations (between the given interferer and T_7) generated by $R_{7\ 3}$:*

T_2	T_6	T_9	T_{11}	T_{12}
1	0	0	0	0
0	2	0	0	0
0	0	0	1	0
0	0	0	0	2
0	0	1	1	0

Thus the first row, for example, denotes:
$(|f_2 - f_7| > 1) \wedge (|f_6 - f_7| > 0) \wedge (|f_9 - f_7| > 0) \wedge (|f_{11} - f_7| > 0) \wedge (|f_{12} - f_7| > 0)$.
The constraint represented by at least one row must be satisfied. Note that the last row is redundant and can be removed. If the constraint represented by the last row is satisfied, then the constraint of the third row is already satisfied.

It would be necessary to merge the constraints generated by all $R_{i\,j}$ and remove redundant constraints. For any moderately large problem it appears that this could only be done selectively. A selective approach would involve adding such non-binary constraints only for receivers where the assignments generated using binary constraints did not give an adequate signal-to-interference ratio.

A constraint free approach might allow the assessment of any potential advantages of the non-binary approach without the need to identify, select and store the non-binary constraints first. Any deficits in the required signal-to-interference ratios become part of the cost function C. When C is reduced to zero the signal-to-interference ratio is adequate at each receiver, and so any possible non-binary (and binary) constraints are satisfied.

Area coverage problems can be created with a benchmark generator which allows transmitters to be located using typical non-uniform distributions. Receiver locations can be generated in two ways. The *Voronoi polygon* surrounding each transmitter location is determined. Such a polygon contains all points closer to the transmitter than to any other transmitter. A small example with most of the Voronoi polygons shown appears in Fig. 4. The *reception points* are the corner points of the Voronoi polygons, and are assumed to be the worst case reception points if the interior of a Voronoi polygon is served by the transmitter that the polygon contains. Additionally, a set of 10,000 *test points* are defined by overlaying the square with a 100 × 100 regular square grid. The reception points are used in the algorithm and the test points are used to evaluate the coverage of the area achieved by the assignment.

A simple propagation model is used for comparison purposes. If reception point r_i, is tuned to

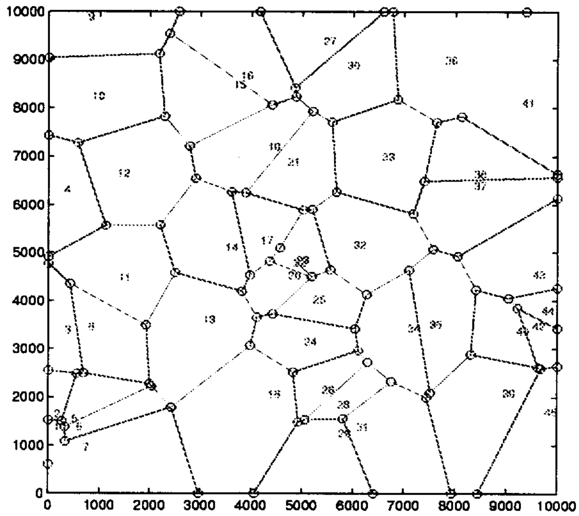


Figure 4: 45 transmitters and 111 reception points were placed as shown. The transmitters are the numbers in the middle of the cells and the reception points are the circles at the cell vertices. Reception points are also placed around the perimeter of the square at all positions equi-distant from two transmitters.

transmitter T_k then the signal strength, S_i, at r_i, is assumed to be given by:

$$S_i = \frac{P_k}{d_{ik}^{\gamma}}$$

where P_k is the power of transmitter T_k, d_{ik} is the distance between transmitter T_k and the receiver r_i and $2 < \gamma \leq 4$. For simplicity, P_k is assumed to be the same for all transmitters.

The total interference I_i, at reception point r_i tuned to transmitter T_k, is given by:

$$I_i = \sum_{\substack{j=1 \\ j \neq k}}^{n} \frac{P_j}{d_{ij}^{\gamma}} \theta$$

where n is the number of transmitters. For each j, θ is assumed to be taken as

$$\left.\begin{array}{l} \theta = 10^{\frac{-\alpha(1+\log_2 f)}{10}} \\ \theta = 1 \end{array}\right\} \text{ if } f \left\{\begin{array}{ll} \neq 0 & \text{(adj. channel)} \\ = 0 & \text{(co-channel)} \end{array}\right.$$

where f is the channel separation between the wanted transmitter T_k and the interfering transmitter T_j. α is an attenuation factor for adjacent channel interference.

A simulated annealing algorithm is applied in the hybrid form described in 3.9, but with a cost function C defined over m reception points by:

$$C = \sum_{i=1}^{m} \left(\frac{S_i}{I_i} - \sigma \right)^a \text{ for all } \frac{S_i}{I_i} < \sigma$$

where a is an even integer and σ denotes the required signal-to-interference ratio. Additionally, the algorithm is modified so that a minimum span can be determined when it is only required that $p\%$ of the test points meet the required signal-to-interference ratio σ.

It is also possible to generate binary constraints for any value of σ as outlined previously. When determining the constraint involving f_k and f_r given by r_i, the expression

$$\sum_{\substack{j=1 \\ j \neq k}}^{n} \frac{P_j}{d_{ij}^{\gamma}} \theta$$

is replaced by

$$\frac{P_r}{d_{ir}^{\gamma}} \theta.$$

These binary constraints can be solved using the algorithms and lower bounding methods in FASOFT. The value of p achieved (evaluated according to the non-binary model) for different values of σ can be determined and σ can be adjusted accordingly.

Before outlining the results, it is important to point out that when the coverage p is 100%, lower bounds can still be determined for the non-binary model. This can be done either for the reception points or for the test points (or for the union of the two sets). The bound will be presented here for the test points. Assume that with a given value of σ the cost function C, evaluated on the test points, is zero. Then for each test point i

$$\frac{S_i}{I_i} \geq \sigma.$$

Thus

$$\sum_{\substack{j=1 \\ j \neq k}}^{n} \frac{P_j}{d_{ij}^{\gamma}} \theta \leq \frac{S_i}{\sigma}$$

and as each term in the summation is non-negative, it follows that for each value of j, $(j \neq k)$

$$\frac{P_j}{d_{ij}^{\gamma}} \theta \leq \frac{S_i}{\sigma}.$$

This simply states that the binary constraint generated by test point i, wanted transmitter T_k and interfering transmitter T_j is satisfied. By using all receiver points and all possible interfering transmitters T_j, it can be seen that all binary constraints are satisfied. The following theorem is then immediate:

Theorem 4 *Let $C(\sigma)$ be a set of binary constraints generated at signal-to-noise ratio σ. Then any lower bound for the span when these binary constraints are satisfied is a lower bound for the span according to the non-binary model (i.e. with $C = 0$) at the same value of σ.*

Although this lower bound appears weak at first sight, it has proved tight for some examples.

Now take α = 15 dB/Octave and a = 2. Applying both the binary constraint method, the non-binary constraint method and the lower bound of Theorem 4 to the example in Fig. 4, the results in Table 1 are obtained. Although

% test points with $SIR \geq \sigma$	95	97	98	99	100
Binary span achieved at 9dB	8	8	8	8	11
Non-binary span achieved at 9dB	6	6	7	8	9
Non-binary lower bound on span (calculated at test points-σ=9dB)					8
Binary span achieved at 17dB	12	14	14	15	20
Non-binary span achieved at 17dB	13	14	17	19	21
Non-binary lower bound on span (calculated at test points-σ=17dB)					13
Binary span achieved at 25dB	24	24	26	26	34
Non-binary span achieved at 25dB	26	28	29	30	39
Non-binary lower bound on span (calculated at test points-σ=25dB)					23

Table 1: Results comparing non-binary and binary methods for the example shown in Fig. 4

these results are not yet complete, results for several networks suggest a number of tentative conclusions:

1. The use of this non-binary "no constraint" approach does lead to somewhat smaller spans, as expected, when the network and σ are both small. For large networks and large values of σ the results are actually worse than using binary constraints. This is because of the time taken to compute the cost function C in the algorithms. Meta-heuristic algorithms

which explore very large search spaces require cost functions which can be quickly evaluated if they are to be effective. In fact, taking the results in Table 1 as typical, the results using binary constraints take only a few minutes, where the results using the non-binary "no constraint" approach take several days to compute.

2. The assignments achieved by the binary constraint method usually meet one of the lower bounds for these binary constraints described in section 4.

3. The lower bound of Theorem 4 is sometimes tight. When it does not appear tight it is unclear whether it is the bound or the (non-binary) span achieved which is weak.

4. For small problems, where the binary span is greater than the non-binary span, most or all of the difference can be recovered by repeatedly identifying the reception points where the signal-to-interference ratio is inadequate, selectively strengthening the binary constraints involved and repeating the binary constraint algorithm.

Thus it seems that a binary constraint algorithm should always be accompanied by an evaluation routine for the assignments obtained, which is used to reconsider and possibly revise the current constraints. It may be that non-binary constraints still have a role, but perhaps they are best added as revisions to the original constraints during this evaluation process.

6 Conclusions

This paper has outlined the problem representations, algorithms and evaluation methods used in frequency assignment. Possibly the most important message is that the effectiveness of the methods and algorithms may be capable of accurate evaluation. In particular, authors presenting new algorithms in the literature should always present them in a way that allows detailed comparison with existing algorithms. This has not always been the case in the past.

7 Acknowledgements

The authors are grateful to Jim Finnie of the U.K. Radiocommunications Agency, Ray Bradbeer of DERA Malvern and Joachim Samuelsson of ComOpt AB for many useful conversations. They also wish to acknowledge help from several others who have provided data for evaluation.

W.J. Watkins is supported by the U.K. Radiocommunications Agency. S.M. Allen is supported by the U.K. Engineering and Physical Sciences Research Council

References

[1] Zoellner, J.A. and Beall, C.L., "A breakthrough in spectrum conserving frequency assignment technology", IEEE Transactions on Electromagnetic Compatibility, EMC-19, 3, August 1977, pp313–319.

[2] Lanfear, T.A., "Graph theory and radio frequency assignment", NATO EMC Analysis Project No. 5, NATO Headquarters, 1110 Brussels, 1989.

[3] Metzger, B.H., "Spectrum management technique", *presented at* "38th National ORSA meeting", Detroit, MI, Fall 1970.

[4] Hurley, S., Smith, D.H. and Thiel, S.U., "FASoft: A system for discrete channel frequency assignment", Radio Science, 32, 5, 1997, pp1921–1939.

[5] Jensen, T.R. and Toft, B., "Graph Coloring Problems", New York, Wiley, 1995 (ISBN 0 471 02865 7).

[6] Hansson, P., Lewer-Allen, C. and Samuelsson, J., "Breakthroughs in frequency planning", Mobile Communications International, Dec.97/Jan.98, pp89–90.

[7] Prosser, P., "Hybrid algorithms for the constraint satisfaction problem", Computational Intelligence, 9, 3, 1993, pp268–299.

[8] Leung, D.S.P., "Application of the partial backtracking technique to the frequency assignment problem", *in* "IEEE Int. Symp. on Electromagnetic Compatibility", 1981, pp70–74.

[9] Hale, W.K., "New spectrum management tools", *in* "Proc. IEEE International Symposium on Electromagnetic Compatibility", IEEE Press, Piscataway, N.J., 1981, pp47–53.

[10] Lochtie G.D. and Mehler, M.J., "Subspace approach to channel assignment in mobile communication networks", IEE Proc.-Commun., 142, 3, June 1995, pp179–185.

[11] Funabiki, N. and Takefuji, Y., "A neural network parallel algorithm for channel assignment problems in cellular radio networks", IEEE Transactions on Vehicular Technology, VT-41, 4, November 1992, pp430–437.

[12] Kim, J-S., Park, S.H., Dowd, P.W. and Nasrabadi, N.M., "Cellular radio channel assignment using a modified Hopfield network", IEEE Transactions on Vehicular Technology, VT-46, 4, November 1997, pp957–967.

[13] Hurley, S., Smith, D.H. and Valenzuela, C., "A permutation based genetic algorithm for minimum span frequency assignment", to appear.

[14] Kirkpatrick, S., Gelatt, C.D. and Vecchi, M.D., "Optimization by simulated annealing", Science, 220, 4598, May 1983, pp671–680.

[15] Costa, D., "On the use of some known methods for T-colouring of graphs", Annals of Operations Research, 41, 1993, pp343–358.

[16] Hurley, S. and Smith, D.H., "Fixed spectrum frequency assignment using natural algorithms", *in* "Genetic Algorithms in Engineering Systems", Institute of Electr. Eng., London, 1995, pp373–378.

[17] Glover, F., "Heuristics for integer programming using surrogate constraints", Decision Sci., 8, 1997, pp3–28.

[18] Glover, F., Taillard, E., and de Werra, D., "A user's guide to tabu search", Annals of Operations Research, 41, 1993, pp3–28.

[19] Dorne, R. and Hao, J-K., "Tabu search for graph coloring, T-colorings and set T-colourings", *in* "Proceedings of 2nd International Conference on Metaheuristics", MIC 97, Sophia-Antipolis, 1997, to appear.

[20] Smith, D.H., Hurley, S. and Thiel, S.U., "Improving heuristics for the frequency assignment problem." European Journal of Operational Research, 107, 1, May 1998, pp76–86.

[21] Allen, S.M., Hurley, S., Smith, D.H. and Thiel, S.U., "Using lower bounds in minimum span frequency assignment", *in* "Proceedings of 2nd International Conference on Metaheuristics", MIC 97, Sophia-Antipolis, 1997, to appear.

[22] Gamst, A., "Some lower bounds for a class of frequency assignment problems", IEEE Transactions on Vehicular Technology, VT-35, 1, February 1986, pp8–14.

[23] Smith, D.H. and Hurley, S., "Bounds for the frequency assignment problem". Discrete Mathematics, 167/168, 1997, pp571–582.

[24] Smith, D.H., Allen, S.M. and Hurley, S., "Lower bounds for channel assignment", *in* "Radio Channel Assignment", (ed. R. Leese), OUP, to appear.

[25] Allen, S.M., Smith, D.H. and Hurley, S., "Lower bounding techniques for frequency assignment", Discrete Mathematics, to appear.

[26] Volgenant, T. and Jonker, R., "A branch and bound algorithm for the symmetric traveling salesman problem based on the 1-tree relaxation", European Journal of Operational Research, 9, 1982, pp83–89.

[27] Volgenant, A., "Symmetric traveling salesman problems", European Journal of Operational Research, 49, 1990, pp153–154.

[28] Prim, R.C.. "Shortest connection networks and some generalizations", Bell System Technical Journal, 36, 1957, pp1389–1401.

[29] Carraghan, R. and Pardalos P.M., "An exact algorithm for the maximum clique problem", Operations Research Letters, 9, November 1990, pp375–382.

[30] Dunkin, N.W. and Jeavons, P.G., "Expressiveness of Binary Constraints for the Frequency Assignment Problem", *in* "DIAL-M Workshop, 3rd Annual ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM'97)", Budapest, October 1997.

PAPER TITLE : **Keynote - Frequency Assignment : Methods and Algorithms**

AUTHOR : D.H. Smith et al

NAME : K.S. Kho

QUESTION :

It is reported that FA Soft achieved for cellular telephone networks: 5% of optimality and for pseudo-military-radio nets : 10%. Radio military tasks are normally more resistant to interference. Could you explain the reasons for the difference in the optimality achieved?

ANSWER :

The estimates quoted apply for a given set of binary constraints. The adequacy (or necessity) of a particular set of binary constraints is a separate issue which we have only addressed for area coverage problems.

NAME : E. Tsang

QUESTION :

Can you elaborate FASoft (what techniques are used in it)? Can you explain how constraint strength can be done automatically?

ANSWER :

FASoft contains a large variety of algorithms, some good, some not very good. The best algorithms for minimum span problems are hybrids of sequential/Calm search (or sequential/SA). The fixed spectrum problems non-hybrid SA and TS algorithms are available.

Automated constraint strengthening can be achieved if the constraint construction, assignment production and evaluation are all included in one routine. The evaluation part must record violating transmitters and pass this information to the constraint production part that will then select the constraints to strengthen.

PAPER TITLE : **Keynote - Frequency Assignment : Methods and Algorithms**

AUTHOR : D.H. Smith et al

NAME : G. Wyman

QUESTION :

What computational resources are required for:

a) assessment of lower hand
b) assignment algorithm

and have any trends emerged from the trials undertaken?

ANSWER :

a) For cellular problems strong bounds can be obtained in seconds. For non-cellular problems it may be necessary to limit the number of paths on which excess variables are defined for the method given to be practical. Run times may still be several hours for problems with thousands of transmitters. The perfect two matching bands is faster and often as strong.

b) The assignment may take seconds, minutes or hours depending on the size and difficulty of the problem. However, most of the computational effort is devoted to final small improvements. Fairly good results are obtained quickly.

NAME : I. White

QUESTION :

1. How do you know how near to optimum a solution is?

2. What are the major unsolved problems in frequency assignment?

ANSWER :

1. Given a set of binary constraints the use of lower bands shows how close to optimum a minimum span solution approaches.

2.
 a) The adequacy of a set of binary constraints remains an open operation.
 b) The determination of strong lower bands for the cost function is fixed spectrum problems.

Application of New Techniques to Military Frequency Assignment.

Ray Bradbeer
PC 308, DERA,
St Andrews Road,
Malvern, Worcestershire,
WR14 3PS, UK.

1. SUMMARY

With increasing use of the RF spectrum, rapid and effective frequency assignment tools are an essential element in the maintenance of effective communications. An abstraction of the VHF Combat Radio Frequency Assignment problem is outlined and discussed in this paper. A key feature of this problem is the wide difference in the magnitude of the frequency separation constraints which arise as a result of radios being located in the same vehicle or site (cosited), and those which apply between nets which do not have cosited members. This feature is used to gain insight into the problem. It is inferred that, within the vast total solution space, it must be assumed that there are a great many optimal or near optimal solutions. A proposed method using graph colouring together with a combinatorial algorithm is outlined. This uses the characteristic large differences between the constraints to focus the search on profitable areas of the solution space.

2. INTRODUCTION

Present day military forces may find themselves needing to move rapidly into an area with little time for familiarisation. Upon arrival, they can expect to find that the radio spectrum is already in use by a whole variety of users, some of whom may not be very well controlled.

It is essential therefore that they have the means to assimilate what knowledge is available concerning the use of the radio spectrum, make plans for their own communications and to change those plans rapidly in response to new information.

The ability to perform rapid and spectrally efficient frequency assignment is a key element of this capability.

The new methods of frequency assignment are showing great promise both for military and civil applications [1].

Clearly it is desirable to develop techniques which show improvements in speed of execution and spectral efficiency. In addition, it may be possible to understand the structure of the solution. Firstly this may enable modifications to be made with minimum disruption. Secondly it may allow the solution process to be sub-divided, to permit sub-sections of a deployment some degree of autonomy without great loss in spectral efficiency.

A comprehensive tutorial paper [1] has already described the techniques of graph colouring, simulated annealing and taboo search, together with other techniques which are applicable to frequency assignment. Readers are referred to that paper for details. The tutorial paper also gives an indication of which methods have been found to work well.

It is proposed here to take an engineering view of the problem, in the hope of exposing, at least to some degree, the basic structure of the problem and hence gain insight into why the good methods work.

It is worth noting at this point that one of the significant conclusions of the CALMA project is that: "Overall there is a strong positive correlation between the amount of problem dependant information used, ... and the quality of the results obtained" [2].

3. PROBLEM ADDRESSED

This paper concentrates on the VHF net radio fixed frequency assignment problem, which illustrates the points particularly clearly. It is believed that the observations apply to other frequency assignment problems but each case requires individual consideration.

The abstraction of the net radio assignment problem is described below.

A basic radio net has a command station and a number of outstations. All the radios operate on the same frequency and when one person speaks, all the others should hear.

It is assumed that the net radio system is serving a military command structure based on headquarters (HQ), shown diagramatically in fig 1. A major headquarters will contain many vehicles and will be connected by communications of many types to other headquarters. Further down the command chain there are a greater number of headquarters but they are smaller; fewer vehicles and fewer transmitter/receiver equipments.

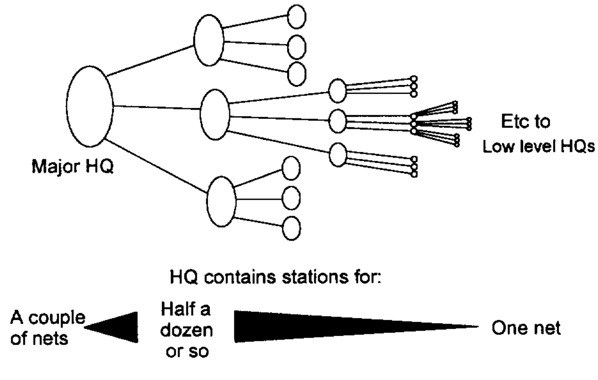


Figure 1: Role of Net Radio

Headquarters containing outstations of nets at one level of command will generally contain the command stations for nets at the next level in the command structure. Major HQs tend to use other types of communications and will contain only one or two net radio command stations. Low level HQs do use net radio extensively, but the nets are well spread out. Each low level HQ will contain stations for only a few nets, perhaps just one working to the next level up and one to the next level down.

The biggest net radio frequency assignment problem is likely to exist at intermediate HQs, which may contain the command stations for half a dozen or so nets, plus the outstations of the nets controlled by the senior level. Stations for a dozen or so nets may be all on the same site.

Generally, the military frequency manager starts with a knowledge of which frequencies can be used and will need to assign these frequencies such that there is no interference between users. If that is not possible, the objective will be to minimise the level of interference and ensure that no user suffers a level of interference such that communications are unworkable. The manager will be particularly keen to ensure that any particularly important connections are interference free. The problem therefore falls into the fixed spectrum category referred to in [1].

4. COSITE INTERFERENCE

When radios are in close proximity, they require much greater frequency separation to work without interference than when they are widely separated geographically. Some indication of the magnitude of this is shown in fig 2. Local interactions between radios in the same vehicle or between radios in the same headquarters give rise to the larger constraints, which are known as cosite constraints. Thus, if we were to draw a graph in which vertices represent nets and edges represent the cosite constraints, we would get a structure which follows the command chain, perhaps with some cross links. In addition, each frequency is subject to many smaller constraints to avoid interference from nets at great distance, which we call farsite.

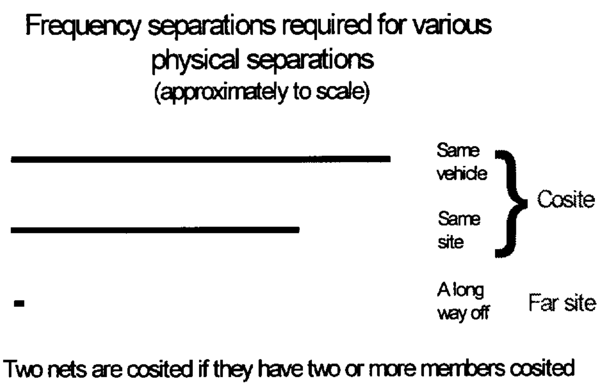


Figure 2: Cosite Problem

5. PROBLEM DIMENSIONS

The military VHF frequency assignment problems typically involve the assignment of hundreds or even thousands of frequencies to a similar number of nets. The total solution space is therefore vast; containing 100s raised to the power of 100s of permutations.

Although not proven, it can be inferred from the simplified frequency assignment problem considered later in this paper that at least this problem does have a great number of dissimilar alternative solutions which are good if not optimal. This is supported by anecdotal evidence. Frequency assignment authorities report that, in general, existing frequency assignment systems either fail to provide a satisfactory solution or are able to generate a number of alternatives.

6. APPLICATION OF COMBINATORIAL ALGORITHMS

Consider now, the application of simulated annealing and genetic algorithms to this problem. Both are based on analogies with nature. It is immediately clear however that the analogies are flawed if the methods are applied in a simple manner directly to the problem.

In the case of simulated annealing. A simple application would take an initial assignment and evaluate the effects of random changes. Accepting or rejecting the change according to the rules which are outlined in the tutorial paper [1]. In real annealing, the atoms of the material, which are all identical, are moved by small excursions until they occupy the correct position in relation to their neighbours. In the case of simulated annealing simply applied to frequency assignment, the frequency constraints applying to each frequency are not identical and the individual frequencies are interrelated in a complex relationship. Thus, a scheme which produces a new solution based on random changes to individual frequencies, without regard to the relative interactions is unlikely to be productive.

The simple application of genetic algorithm is also flawed. Such a simple application might involve the generation of an initial population, from which the individuals giving rise to the lowest overall interference would be selected. These would then be crossed in some way to produce the members of a new generation. If the problem has a number of similarly good but entirely unrelated solutions and the crossing mechanism is applied to two members of the population which are deemed to have relatively high fitness because they are both in the neighbourhood of the same solution, the resulting offspring have a significant probability of showing an improvement compared with their parents. If, upon the other hand, the parents are from entirely different neighbourhoods, the offspring may lie anywhere in the solution space and have a low probability of showing high fitness. In nature, evolution takes place within species. Non viable crosses between entirely different organisms are not even attempted. The concept of speciation has been addressed [3,4,] and the possibility of a generic methods which are able to operate with problems which have multiple global maxima has been addressed [4].

As discussed later in this paper, there is at least a possibility that practical problems are so highly multi-modal that the number of optima available exceeds a practical population size by orders of magnitude.

Goldberg [3] discusses the need to ensure that the genetic algorithm does not concentrate all its population on a single peak of a multi-modal problem. Also the need to avoid the production of poor offspring (lethals) as a result of interbreeding between different species.

In the case of this frequency assignment problem, it is also thought to be necessary to ensure that the number of peaks investigated is limited such that each has a viable population.

These comments suggest that, if either simulated annealing or genetic algorithms are to be applied to this frequency assignment problem, and achieve practical results, either:

a) productive areas of the solution space will need to be identified and the algorithm guided so as to avoid exploration of unproductive areas of the solution space, or

b) the techniques must be applied indirectly in such a way that the problem presented to the combinatorial algorithm has a smaller solution space and is less multi-modal.

As an example of the latter technique, Hurley et al [5] have studied a method in which a genetic algorithm is used to determine the ordering of requirements presented to a sequential (greedy) algorithm.

To achieve the former, some insight into the key features of the different solution neighbourhoods is required.

7. IDENTIFYING KEY FEATURES

The key is believed to be the cosite constraints. The difference between the size of the constraints is particularly marked in the VHF net radio assignment problem considered here.

It is immediately clear that, if there are a number of frequencies which need to be mutually separated by a large cosite separation, they must be spaced out down the available spectrum in a sensible manner. The basis of the proposal is to use this to introduce some problem specific information.

8. PROBLEM SIMPLIFICATION

If, initially, we simplify the problem by assuming that there are just two types of constraint; cosite and farsite, rather than different degrees of each. Further, assume that the channel separation is such that the farsite constraint will be met if nets between which this constraint applies are assigned different channels. If we also assume that the available frequency space consists of adjacent channels, with the lowest adjacent to the highest, so that there are no frequencies with any special properties with regard to those adjacent; ie, a circular frequency space is assumed.

9. SIMPLIFIED PROCEDURE

It is now possible to construct a graph (G_c) using only the cosite constraint information. In this graph, vertices represent nets and edges represent cosite constraints between nets. Thus the vertices representing a pair of nets, will be joined by an edge if they have at least one pair of cosited stations.

This graph can be expected to contain a highly connected set of vertices representing the nets which have stations in the intermediate level of headquarters. There will be many vertices representing the nets lower in the command hierarchy, but these will be relatively sparsely connected. It is believed that colouring this simplified graph using the minimum number of colours will generally be practical. Thus far it has been shown that simple backtrack algorithms are not practical, however, successful colouring of a small number of large problems has been achieved using the algorithm due to Trick [6]. In any case it is worth pursuing the argument as it gives further insight into the problem.

The process for a very small deployment is shown in figure 3.

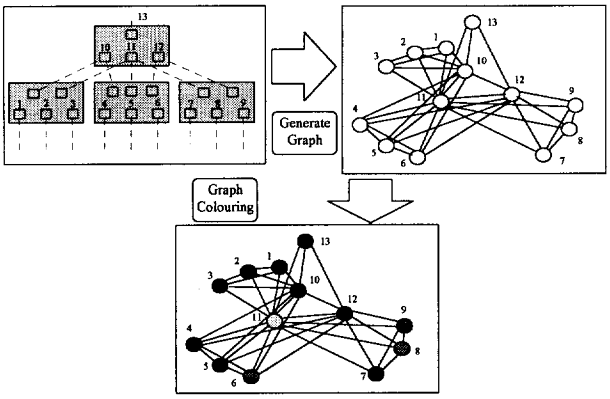


Figure 3: Derivation of cosite graph

Assuming that the graph colouring has achieved a colouring using the minimum number of colours, this will yield a chromatic number $\chi(G_c)$.

Suppose the available frequency band is now divided into equal sub-bands. If the frequency span is wide enough to accommodate the requirements, there will be a *central section* at the centre of each sub-band which is separated from the similar section in the next band by the cosite constraint. There will be *side sections* on either side of the central section. This is shown in figure 4.

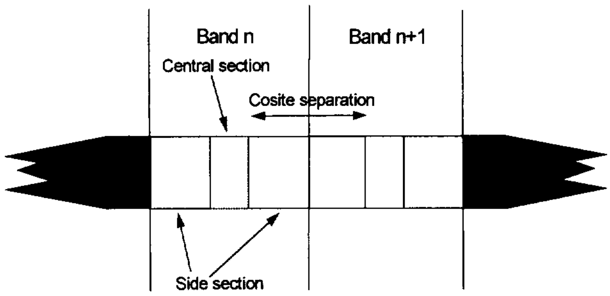


Figure 4: Sub-band Structure

For the simplified procedure it is possible to envisage making a random assignment of colours to the frequency sub-bands, thereby assigning each net to one of the sub-bands.

The nets can then be categorised according to which section of the sub band they need to occupy. Nets which require a frequency from the central section, have been defined as *restricted.* Those which can be near one side of the sub-band but not the other, have been defined as *semi-restricted,* either *upper* or *lower* according to which part of the band can be occupied. Those which can occupy any frequency within the sub-band, have been defined as *not restricted.* This categorisation can be derived from the graph. If the vertex representing a particular net is connected by edges to other vertices taking the colours of both the neighbouring bands in the frequency space, it is categorised as *restricted.* If the vertex is connected to one of the neighbouring colours but not the other, it is *semi-restricted, upper* or *lower.* Otherwise it is *not restricted.* This categorisation is illustrated in fig 5.

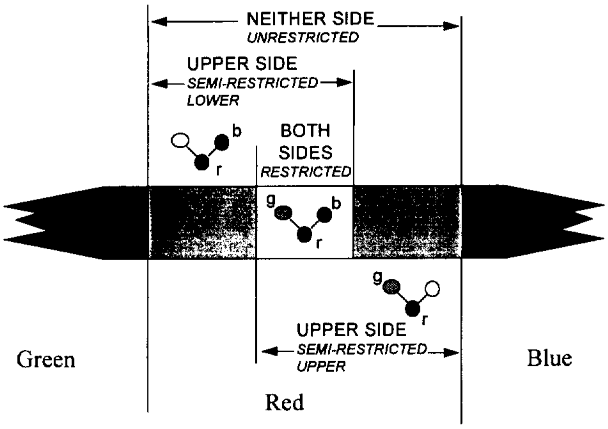


Figure 5: Sub-band Assignment

If individual frequencies can now be assigned within their allotted sections, in such a way that the far site constraints are met, then an assignment which meets all the constraints will have been achieved. However, if this process fails, there are further freedoms which can be exploited.

We might try to improve the solution provided by this method by applying a combinatorial algorithm. To be productive, this should really work within the constraints implied by the structure laid down by the assignment of colours to bands. Nevertheless there are some freedoms which are not exploited by the basic process and clearly could be explored, for example:

a) net assignments can move within the portion of the spectrum to which they have been assigned, according to whether they are restricted, semi-restricted or not restricted. This freedom can be used to avoid additional incompatibilities due to, for example, intermodulation products and spurious responses/emissions.

b) Assignment within the portion of the spectrum as in a) above guarantees that the criteria will be met. In general, further constraining some of the frequencies will give additional freedom to others. Thus the assignment of a net to a section of the available spectrum is only a guide to its final assignment. Some assignments will be able to move outside their assigned section.

c) It will be possible to recolour some vertices of the graph without affecting the rest of the graph. Clearly the nets which these vertices represent can move to any band assigned a colour which can be adopted by the associated vertex. In principle, this can be extended to the re-colouring of sub-graphs rather than individual vertices. Note that when a net is moved from one band to another it must be recategorised in respect of the level of restriction applying.

Thus it is possible to envisage a procedure following the initial assignment, in which a combinatorial algorithm is applied. Several strategies could be applied, for example, the combinatorial algorithm might first position all the restricted frequencies, some of which may need to be placed outside the central sections. Semi-restricted and finally unrestricted frequencies would then be placed using the combinatorial algorithm.

However, this is not a practical algorithm because the simplifying assumptions will not, in general, apply. Nevertheless, it is possible to use the simplified case to give an indication of the degree to which the problem is multi-modal.

10. MODALITY OF THE PROBLEM

Notice that there will be at least one node in the graph which is surrounded by all the other colours. If this were not so for a particular colour, it would be possible to recolour every vertex of that colour to at least one other colour and thereby reduce the number of colours used. Thus there will be at least $\chi(G_c)$ restricted vertices whatever happens in the graph colouring or the assignment of colours to bands. Any solution which satisfies the constraints must space these frequencies in the available spectrum such that the cosite constraints are met.

This requirement is sufficient to eliminate large tracts of the solution space. Nevertheless, there are many ways of placing these *framework frequencies*. In the simplified case we can imagine them being placed in the central section of the frequency sub bands. In the absence of problem specific evidence to the contrary, it has been assumed that any ordering of the colours is equally likely to yield a solution. There are $(\chi(G_c) - 1)!$ such orderings, assuming that the same ordering in a different orientation on the circular frequency space are considered to be the same solution.

Note that the very small network considered in fig 3 has a chromatic number of 6. In practice, chromatic numbers of 10 to 20 are expected. There is reason to believe therefore that, in the case of a realistic problem subject to the simplifying assumptions mentioned above, there may be between 10^5 and 10^{17} distinctly different solutions, or species in the genetic algorithm sense.

Real problems, without the simplifying assumptions, need to exploit the structure of the allotment as well as the net laydown, so they probably have a smaller number of distinct solutions, although this is difficult to estimate. Nevertheless, it is considered that, unless it can be shown otherwise, we must assume that practical cases also have a large number of distinct solutions. It is interesting to note that Hurley et al examined the entire solution space for a 12 transmitter problem and found 4×10^6 permutations giving an optimum span, from a total of 479,001,600 [5].

11. PRACTICAL CASE

The major deficiency in the simplified case is the assumption that the frequency allotment is contiguous.

To overcome this problem, some means is required to divide the available allotment into $\chi(G_c)$ bands such that the number of frequencies available in each category (restricted, semi-restricted etc) matches the number required for each colour band.

A combinatorial algorithm can be applied to this problem; a further example of the indirect application of combinatorial algorithms.

One technique under examination is to place $\chi(G)$ - 1 movable masks of width equal to the cosite separation onto the allotment. Each central band is defined by the region between the masks. The band boundaries are under the masks, as shown in fig 6.

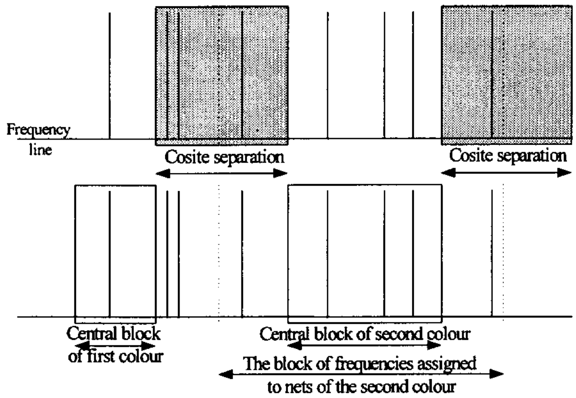


Figure 6: Finding Block Boundaries

The process has the freedom to slide the masks along, move the block boundaries under the masks, and change the order of the assignment of colours to bands.

Having obtained the best match using this process, the recolouring of vertices is applied to give an improvement if possible.

The assignment is then attempted, using a combinatorial algorithm to assign in order of difficulty as outlined above.

Finally a combinatorial algorithm can be applied to the whole solution if required.

12. CONCLUSION

The military combat radio frequency assignment problem has been outlined.

This particular problem can involve the assignment of 100s or even thousands of nets using an allotment of a similar number of frequencies. Giving rise to a very large total solution space.

The problem is characterised by constraints which fall into two groups, cosite and farsite with a large ratio between them.

It has been shown that the problem is likely to demonstrate highly multi-modal behaviour.

A proposed algorithm which takes account of these features has been outlined.

REFERENCES

[1] D.H. Smith, S.M. Allen, S. Hurley and W.J. Watkins, Frequency Assignment Methods and Algorithms, Proceedings of NATO Research and Technology Organization, Sensors and Electronics Technology Panel Symposium on Frequency Assignment, Sharing and Conservation in Systems (Aerospace), October 1998, Aalborg, Denmark.

[2] S. Tiourine, C. Hurkens, J.K. Lenstra, An Overview of Algorithmic Approaches to Frequency Assignment Problems, Eindhoven University of Technology, August 1995.

[3] D.E. Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley, 1989.

[4] A. Pétrowski, An Efficient Hierarchical Clustering Technique for Speciation, Third Conference on Artificial Evolution, AE97, Nimes, France, October 1997.

[5] S. Hurley, D.H. Smith and C Valenzuela,
A Permutation Based Genetic Algorithm for Minimum Span Frequency Assignment, to appear.

[6] COLOR.C: Easy code for graph coloring, Michael A Trick, Carnegie Mellon University.

PAPER TITLE : **Paper 1 - Application of New Techniques to Military Frequency Assignment**

AUTHOR : R. Bradbeer

NAME : K.S. Kho

QUESTION :

You focus on new technology for frequency assignment. However, the limitation of the deployed CNR frequency management assignment system lies on the physical limit of the span of the CNR. VHF band with respect to cosite requirement (numbers of radios cosited etc.?) Could you comment on this?

ANSWER :

It is assumed that Combat Radio Co-Siting will continue to be required. It is hoped that the techniques suggested will provide a better indication of the origin of frequency assignment problems in specific cases. For example, lack of sufficient span to accommodate co-site constraints can be detected and reported to the Frequency Manager.

NAME : M. Vant

QUESTION :

Could your technique be adapted to the situation in which there is jamming on the battlefield and some frequencies are not available?

ANSWER :

It is hoped that algorithms based on Combinational Algorithms and Graph Colouring can be used to improve responses to changes in the availability of Spectrum. Jamming is clearly one reason why this availability could change.

SPECTRUM CERTIFICATION - THE FIRST STEP

Stanley R. Green
Joint Spectrum Center /IIT Research Institute
185 Admiral Cochrane Dr
Annapolis, Maryland 21401 USA

Craig A. Scammon
Department of Commerce/NTIA
1401 Constitution Ave, NW
Washington, D.C. 20230 USA

ABSTRACT

Assigning frequencies for electronic systems on the ever diminishing electromagnetic spectrum is a difficult process, compounded by the different modeling approaches that employ complex frequency assignment algorithms. These algorithms are based on the available equipment parameters and environmental data. The underlying assumption is that the data being used is the best and the most accurate data available. This assumption is rarely if ever correct. But with today's highly sophisticated technology in electronic equipment, having the most accurate data available for use in spectrum management systems such as the Joint Spectrum Management System for Windows ($JSMS_w$) frequency assignment model is critical.

Collecting accurate data begins with the request for spectrum allocation support, via the application for electronic equipment certification. This data collection continues through equipment design, procurement, and operational deployment of the electronic system. Traditionally, this collection of data began by entering the pertinent system data on a paper form. Today, a template for this form has been created, and the process for entering data has been automated by using the Spectrum Certification System (SCS).

The National Telecommunications and Information Administration (NTIA) has been working on a new approach to capture the required data, using a Smart Interface Diagram (SID) technology. A computer software program called the Equipment Location - SID (EL_SID) that will automate the SID is under development. This program provides a graphical, icon-based user interface supported by sophisticated logic that captures inter- and intra system relationships and prompts the applicant to enter minimal but pertinent system parameters. The EL_SID interface will simplify the collection of data, enabling the applicant to enter the most comprehensive, and the most accurate, information available for particular operational characteristics of the electronic equipment. The EL_SID interface will also enhance earlier efforts at defining system characteristics by identifying the actual relationship between equipment parameters for the links in a network, and thus will provide the best available operating characteristics.

This paper will describe the spectrum allocation process, provide a historical background of data entry, and look to the future for spectrum allocation and assignment.

BACKGROUND

In planning telecommunication systems, the United States Government agencies are required to develop systems for operational use in

Paper presented at the RTO IST Symposium on "Frequency Assignment, Sharing and Conservation in Systems (Aerospace)", held in Aalborg, Denmark, 5-7 October 1998, and published in RTO MP-13.

accordance with the applicable portions of the National Tables of Frequency Allocation and the provisions of the *NTIA Manual* [1] which includes technical standards and policy issues concerning the use of the radio frequency spectrum.

Government agencies planning the use of, conducting experiments relating to, or developing and obtaining telecommunication systems requiring the use of radio frequencies are required to take all reasonable measures to ensure that such systems when placed in their intended operational environments will neither cause nor be subject to harmful interference to, or from, other authorized users.

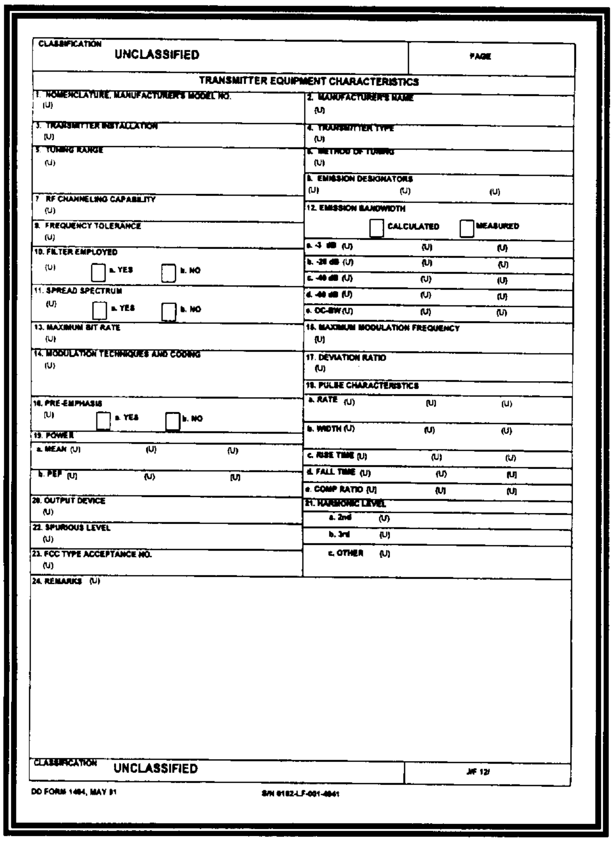
To assist government agencies in meeting the above responsibilities, NTIA established the Spectrum Certification Process to ensure that radar and communications equipment operating within an environment meet certain rules, regulations, and constraints consistent with the particular piece of equipment. All major systems that transmit or receive radio signals must be certified. Certification ensures that critical information about the equipment is provided so the equipment can either be protected from harmful interference from other spectrum users or permitted to operate in a particular frequency band without causing harmful interference to other users of the radio spectrum. The Spectrum Certification Process enables the Spectrum Planning Subcommittee (SPS) to review government telecommunication systems and subsystems. These reviews are ongoing during the four stages of system development (**Stage 1 - Conceptual, Stage 2 - Experimental**, **Stage 3- Developmental**, and **Stage 4 - Operational**), before a frequency assignment is granted. During the first three stages, the SPS provides guidance to the agency to ensure that the development of the system or subsystem will meet NTIA regulations when the Stage 4 operational certification is requested. Certification of spectrum support for telecommunication systems or subsystems at Stage 4 is a prerequisite for NTIA authorization of frequency use from a station with an operational station class (i.e., other than experimental). The Stage 4 certification provides restrictions on the operation of the system or subsystem as may be necessary to prevent harmful interference.

In addition to reviewing the development of new telecommunication systems, modifications of the technical or operational characteristics of an *existing* systems or subsystem that could have significant impact on the radio frequency spectrum must also be submitted to the SPS for review. During the SPS's review, consideration is given to system compliance with prevailing spectrum management policy, allocations, regulations, and technical standards (Government, National, and International), and the predicted degree of electromagnetic compatibility (EMC) between the proposed system and the electromagnetic environment. Upon assessment of a proposed system or subsystem, considering these criteria and any other pertinent factors, the SPS will provide recommendations, along with supporting documentation, to NTIA.

After the criteria for systems have been met, the Spectrum Certification Certificate will then be issued by NTIA. Authorization for frequency use will not be granted for a system until notice is received that Stage 4 frequency support for the system has been certified. Authorization for frequency use will be granted for the system, subject to any limitations or constraints contained in the Spectrum Certification Certificate.

DD FORM 1494 FOR FREQUENCY ALLOCATION

The NTIA developed forms 33, 34, and 35 for use by the civil sector. The DD Form 1494-Application for Frequency Allocation was developed by the Department of Defense. These forms are the original format used for the submission of data. The civil and military forms are similar in content and NTIA accepts both



DD Form 1494

formats for review.

NTIA reviews the DD Form 1494 for certification. When correctly filled out, this form provides technical information to permit officials to evaluate the request for use of electronic spectrum frequency(ies).

With more sophisticated, complex systems emerging, however, the paper copy of the DD Form 1494 no longer provides a means to enter details about the systems' operating characteristics. As the technology improved, the equipment functions became more complex, and often the application would exceed 30 pages of technical information, and because of the limited size of the form itself, information would be truncated or deleted that would be helpful in processing the request.

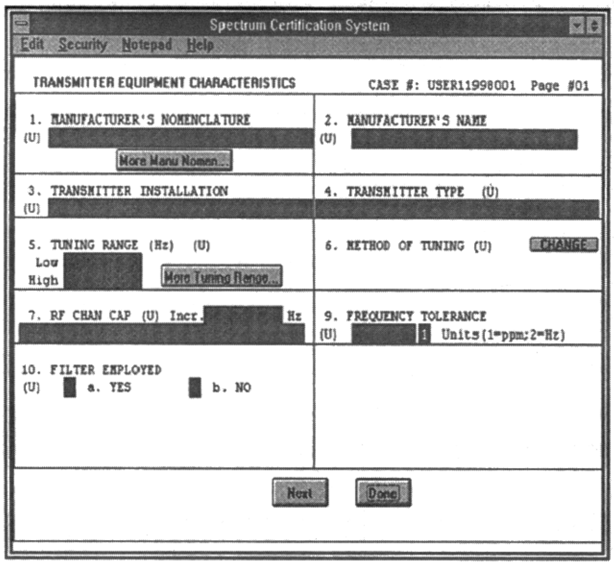
Further, the form did not facilitate the need for information concerning the operational relationships between data elements. Additional data or information concerning the relationships between the data elements was presented in the general remarks section, which is an unstructured text field. Information provided in these text fields was almost impossible to use, or to extract data from, in automated analysis algorithms. As the technology advanced, applicants no longer understood what type of information was required nor how to present that information. For example, applicants did not know which data elements were required for a particular system or which data fields did not apply to a particular system.

SPECTRUM CERTIFICATION SYSTEM

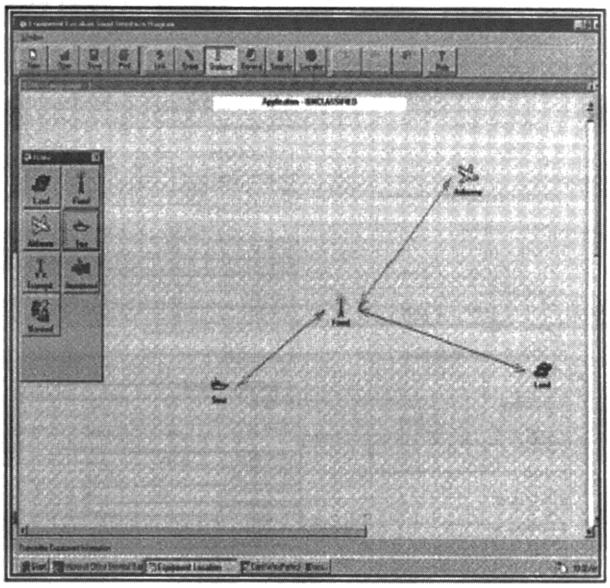
In the 1990s, the Joint Spectrum Center developed an automated system, the Spectrum Certification System (SCS), to enhance the Spectrum Certification Process. A template for the DD Form 1494 has been created, and the process for entering and storing data has been automated by using the SCS. The automated database, populated with these forms can be queried for various type of information. In addition, the SCS contains a diagram that can be used to further clarify the use of the equipment employed. To ensure consistency of the data wherever possible, data checking is performed.

As part of the certification review process, technical parameters must be compared with NTIA's technical standards, and an EMC analysis must be conducted to determine if there is a possibility of potential interference between the proposed system and systems in the environment. To perform an EMC analysis, specific relationships between data must be known. For example, if a system has several power levels and several different emission bandwidths, NTIA needs to know what power levels are associated with which emission bandwidths to perform an accurate EMC analysis.

Without this input, a so called "worst case" EMC analysis results. The analysis performed is, as such, very conservative. The consequence is that during the SPS Certification Review



Template Created for the DD Form 1494 for Frequency Allocation



Smart Interface Diagram (SID)

Process, limitations or restrictions are imposed on the deployment of the system. Ultimately, a frequency assignment may be denied because an electronic system is predicted to cause interference, simply because relational characteristics have been identified that may not actually exist in the operational equipment.

SMART INTERFACE DIAGRAM

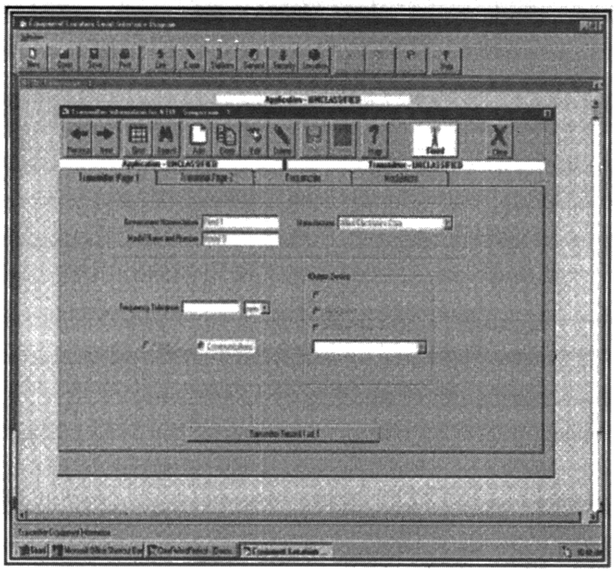
The need to capture information regarding the more complex communications-electronic (CE) equipment and their inter- and intra- systems relationships is becoming more critical, as is the need to perform EMC analyzes that will enable these systems to be certified for frequency assignments.

In lieu of this, NTIA has not only began to redesign the spectrum certification process but also has developed a new technology for capturing and reporting technical data for spectrum certification. This new technology is called the **Smart Interface Diagram (SID)**. SID provides a graphical, icon-based interface supported by sophisticated logic that captures inter- and intra- system relationships and that prompts for applicant to enter, only those system parameters required for the particular system being described.

To fulfill its mission into the 21st century, NTIA is currently creating a new data dictionary to store the data needed and to automate the frequency assignment process. This data dictionary would contain a multitude of complex scenarios that were previously difficult to describe.

A complex communications site might consist of multiple antennas, each with one or more emissions and frequencies transmitting to various receivers. Thus, the data dictionary would provide for the entry of all the possible communications modes at a site, along with the specification of which mode(s) is received at each of its receivers. Specifying only the actual communications modes used, along a particular communications link, eliminates trying to analyze modes that do not exist. A communications scenario becomes more complex when two or more of these sites are linked, as in a network. In which case, a more sophisticated method is needed such as using an interactive diagram. The interactive diagram, SID, is used as an entry level, from which increasing levels of detail are revealed

This diagram contains icons that represent communications equipment. The applicant can drag or drop these icons to any location on the



Smart Interface Diagram Data Entry

computer screen. All possible transmissions and/or receiver modes for that equipment are revealed when the applicant clicks on the corresponding icon. Lines are used to link each transmitter to its intended receiver. An arrow at the end of each line depicts the flow of communication from the transmitter to the receiver. When the applicant clicks on the line, all the possible transmission and receiver modes are revealed that are actually used for a particular transmitter/receiver pair. More than one communications mode can be specified for a link.[2]

Thus, entry and display of complex spectrum certification data is simplified: the applicant simply specifies the communication links and network/system configurations, and is then prompted for a minimum of information based on previous data entered.

The applicant enters the modes of operation that the equipment is designed to use. This information consists of power, frequency, emission bandwidth, and modulation codes. This information can be used to establish link information that specifically details which of the modes the equipment will be using, such as a particular frequency or frequency band, power, emission bandwidth, and modulation codes. A piece of equipment could use 10 different mode combinations, but the applicant is only identifying the use of a subset of the modes.

When the application is forwarded for certification, officials examine the link information and determine if the application can be accepted. The certification officials may impose some constraints for use of the equipment such as a reduced power, or deny a particular modulation code, or restrict the equipment to a particular location. This information is then stored in the data record as the certified data. This information is critical since the equipment could be operating at a reduced capability to be certified.

The database is being designed to support automated interference analysis algorithms. Each data element has a unique field in the database. The database will also have structured data fields that will hold the data containing limitations or restrictions placed on the system.

The EL_SID (database and user interfaces) will collect and store all the technical data (and operational relationships between this data) that accurately describe the system and any restrictions or deployment restraints placed on this system. This data can then be accessed by automated interference frequency analysis routines.

The EL_SID is efficient for obtaining large amounts of data. The system has a sophisticated interface design that will prompt the applicant for the necessary data elements that apply to his particular system and will not prompt him to supply data elements that are not used for his particular system. The user is provided a method of showing the relationship data information that only requires a point and click with the mouse from a provided list of possible relationships.

FREQUENCY ASSIGNMENTS VIA THE SMART INTERFACE DIAGRAM

In an era of diminishing spectrum resources and increasing demands, using the most accurate information is critical. During the frequency assignment process, a mathematical algorithm is used to select suitable frequencies to assign a request. In the past, the equipment parameters used were derived from the DD Form 1494. This information now reflects the equipment characteristics the system actually has been approved to use.

Using the EL_SID interface, applicants will be able to enter the most accurate operational characteristics in order to model the environment and to perform a more precise EMC analysis, rather than an analysis based on "worst case" conditions such as , the highest power or highest antenna gain, the widest emission bandwidth, or most sensitive receiver sensitivity. This new interface will enhance the every stage of the spectrum certification process.

REFERENCES

[1] *NTIA Manual of Regulations and Procedures for Federal Radio Frequency Management*, NTIA, Washington, DC, 1995.

[2] NTIA memo, subject: The Smart Interface Diagram Approach to Frequency Management, June 1996.

PAPER TITLE : **Paper 2 - Spectrum Certification – The First Step**

AUTHOR : S. Green, C. Scammon

NAME : M. Elliott

QUESTION :

How will users have access to the new system and will the US Frequency Management Offices around the world be able to access the database to support their frequency assignment activities.

ANSWER :

The new system will be a stand-alone system available through the US Federal Government. Possible sources in the future could include a web site.

US frequency managers will have access to the database.

NAME : D. Jaeger

QUESTION :

1. Is there any feed-back procedure about practical problems (if they exist) after a certification process?

2. Are possible inter-modulation effects and harmonics, which can not be suppressed to zero, considered in the frequency certification process, too?

ANSWER :

1. Practical problems may be reported back to the Certification agency. Problems which exist may also be reported back to the agency in charge of the frequency assignment.

2. The NTIA manual lists the rules and regulations for each band of operation.

PAPER TITLE : **Paper 2 - Spectrum Certification – The First Step**

AUTHOR : S. Green, C. Scammon

NAME : G. Wyman

QUESTION :

How do you map the information contained in the database to constraints applicable to the assignment algorithm?

ANSWER :

Each critical parameter is given a discrete entry in the database. These parameters are addressable by automated frequency assignment routines.

NAME : D. Jaeger

QUESTION :

Is there any co-ordination between the output power of a e.g. radio transmitter and the EM-environment defined as a certification environment for commercial aircraft?

ANSWER :

The spectrum certification process gives the approval to operate. Enforcement of the correct operation is the function of a different group.

PAPER TITLE : **Paper 2 - Spectrum Certification – The First Step**

AUTHOR : S. Green. C. Scammon

NAME : K.S. Kho

QUESTION :

1. The spectrum Certification System is implemented with JSMS. Is this also the system implemented in the US Navy ASPECTS programme?

2. We are familiar with the form 1494. What other parameters are added to the parameters in this form 1494. Do we need it? If yes, do we have the manpower to process it?

ANSWER :

1. ASPECTS has a certification system but it is not the system implemented in JSMSw The JSMSw produces the DD 1494 which is the official form for the Spectrum certification requests.

2. Additional parameters will be used to describe the emerging technologies (i.e. Spread Spectrum, Frequency Hopping).

SPECTRUM MANAGEMENT USING $JSMS_W$

Thomas C. Hensler
Joint Spectrum Center/IIT Research Institute
185 Admiral Cochrane Dr
Annapolis, Maryland 21401 USA

ABSTRACT

The Joint Spectrum Management System for Windows ($JSMS_W$) is a set of frequency management tools that runs on a personal computer. $JSMS_W$ provides the spectrum manager with the capability to create assignment proposals, edit them and perform validation checking of the proposal. The user then uses $JSMS_W$ to format the proposal for transmission to the US National Authority for approval. Other tools included are an interference reporting capability and an allotment plan generator which enables users to define the frequency resources for the frequency nomination capability.

This paper focuses on the eight discrete steps performed during the $JSMS_W$ frequency nomination process. These steps are to: specify the parameters, select the environment records, create analysis records, cull environmental records not likely to interfere or to be subject to interference from the proposal, compute received interference power and system noise power levels, determine interference-free frequencies, and, then last, rank the proposed frequencies.

BACKGROUND

The original automated Joint Spectrum Management System (JSMS) was developed for the US Department of Defense (DoD) by the Sentel Corporation in 1992. IITRI converted the DOS version of JSMS to Windows and now the program is referred to as JSMS for Windows ($JSMS_W$). A number of $JSMS_W$ versions have been released in recent years; the latest being Version 3.0 which was released at the end of FY98. Version 3.0 is a year-2000 compliant software system that can be used on Windows 3.1, Windows95/98 or Windows NT systems on PC platforms. $JSMS_W$ is written primarily in FoxPro, but also contains C, Visual Basic, FORTRAN, and CLIPS code.

In recent years, the United States National Telecommunications and Information Administration (NTIA) has joined forces with the DoD Joint Spectrum Center (JSC) to create a system that can be used by both US federal government and DoD organizations. Thus, $JSMS_W$ can be used by any organization that is required to forward frequency assignments and proposals for approval and subsequent entry into the US Government Master File (GMF). The $JSMS_W$ editor is tailored to "type of user," so users can work in either the DoD Standard Frequency Action Format (SFAF) or the GMF mode. There are approximately 700 copies of $JSMS_W$ being distributed to DoD and Federal users.

$JSMS_W$ is designed to provide user assistance in the spectrum certification and frequency assignment portions of spectrum management. $JSMS_W$ contains modules to assist in creating frequency assignment and spectrum certification applications, assignment proposal editing and validation, and the analysis of the spectrum environment to identify available frequency channels and potential interference problems. $JSMS_W$ can be used to nominate frequencies and to rank them based on efficient spectrum use. The applicants frequency assignment proposals

Paper presented at the RTO IST Symposium on "Frequency Assignment, Sharing and Conservation in Systems (Aerospace)", held in Aalborg, Denmark, 5-7 October 1998, and published in RTO MP-13.

can be formatted for submission to the NTIA Frequency Management and Records System (FMRS).

SYSTEM DESCRIPTION

$JSMS_W$ consists of 10 major modules listed and briefly described below.

1. **Frequency Assignment (FA)** - This module is used to view, create, edit, import and export assignments and proposals. This module also includes the capabilities to nominate candidate frequencies for proposals and to validate the proposal data before submission into the FMRS.

2. **Spectrum Certification** - This module is used to create applications for electronic equipment certification. A package of analysis tools is included in this module.

3. **Joint Restricted Frequency List (JRFL) Editor** - This module is used only by DoD personnel to enter frequencies that will be considered in interference and Electronic Warfare (EW) analyses. The capability to enter Communications-Electronics Operating Instructions (CEOI) frequency data is included in this module.

4. **Interference Report** - This module is used to document interference problems. These can then be electronically submitted later for resolution.

5. **Interference Analysis** - This module, which uses the same algorithms as the FA module, can be used to analyze a specific interference situation.

6. **EW Deconfliction** - This module is used to analyze situations where electronic frequency jammers can be used.

7. **Engineering Tools** - This module is a collection of analysis tools that include: a link analysis capability with a plot of the terrain between the transmitter and receiver, an HF analysis capability, coverage plots that include power-density calculations with overlaid feature data (e.g., roads and airports), a capability to calculate satellite look angles, a spectrum occupancy plotting capability, and a coordinate conversion capability. Within this module there is also a direct link to NTIA's Minicomputer System Analysis Models (MSAM) that contains a number of additional analysis capabilities, such as propagation models and an intermodulation analysis model.

8. **Allotment Plan Generator** - This module is used to develop frequency plans to be used for deployments. Pre-set allotment plans are included from Chapter 4 of the *NTIA Manual*[1]. In addition to the pre-set plans, users can enter allotment plans that can be used in the frequency nomination process to specify the candidate frequencies available for assignment. A frequency scheduler, included in this module, can be used by a base frequency coordinator to assign frequency use on a daily basis.

9. **System Manager** - This module contains utilities to update standard tables and to perform general housekeeping tasks, e.g., file maintenance.

10. **Compliance** - This module is used

to validate assignments as is done in the FA module. This stand-alone capability is identical to the FA module, but the assignment data does not have to be in the $JSMS_W$ database.

The paper focuses primarily on the Frequency Assignment module and the interference analysis performed during the nomination process.

THE $JSMS_W$ FREQUENCY NOMINATION PROCESS

The $JSMS_W$ nomination process is performed via the eight discrete steps listed below.

Step 1 - Specify parameters for the proposal to be assigned.
Step 2 - Select potential interferer/victim assignments from the environment database.
Step 3 - Create analysis records for the assignments (see Step 2 above).
Step 4 - Cull potential interferer/victim records.
Step 5 - Compute the received interference power level.
Step 6 - Compute the system noise power level.
Step 7 - Determine the predicted interference-free frequencies.
Step 8 - Rank the frequencies (see Step 7 above).

A detailed discussion of these eight steps (described in Reference 2) follows.

Step 1 - Specify Parameters

The user must provide the following parameters (data elements) for the proposal to be assigned: latitude and longitude, power, antenna gain, antenna azimuth, antenna polarization, antenna height, emission designator and radius of mobility. $JSMS_W$ is designed to use default values for all of the above data elements except latitude and longitude. The default values specified in $JSMS_W$ are a function of frequency and representative of the values found in the environment.

The user must also specify the range of candidate frequencies to be tested and either the equipment-tuning increment or the channelization for the frequency band. Using $JSMS_W$, the user can use either the world-wide allocation tables or an allotment plan to specify the candidate frequencies.

Step 2 - Select Environment Records

After the user has provided the necessary parameters, $JSMS_W$ selects for possible interferer transmitter and victim receivers, relative to the proposed transmitter and receiver, based on the range of the candidate frequencies. The collection of environment records include a wider range of frequencies than the range of candidate frequencies to ensure that wide bandwidth transmitters, e.g., television stations, are included in the analysis.

Step 3 - Create Analysis Records

For each environmental record selected (see Step 2), one or more transmitter and receiver analysis records are created. The general rule is that a transmitter analysis record is created for each instance of station class (the type of assignment, e.g., fixed, mobile) power, and emission. Similarly, a receiver analysis record is created for each receiver defined by the environment record. $JSMS_W$ creates the analysis records the first time an interference analysis is performed for a given frequency range. This procedure reduces execution time for subsequent analyses.

Step 4 - Cull Environment Records

$JSMS_W$ eliminates records from the analysis environmental records as soon as it can be determined that no potential interference situation exists. Thus, record culling occurs a number of times in an effort to minimize execution time. The first cull (described in Step 2) is based on frequency. The second cull, performed during this step, Step 4, is based on a radio line-of-sight calculation that uses the standard 4/3 earth radio horizon formula, d_{cull}, in km:

$$d_{cull} = 1.609\,[(2h_{Tx})^{0.5} + (2h_{cull})^{0.5}] \quad \text{(Equation 1)}$$

where

h_{Tx} = height of proposed station antenna, in feet

h_{cull} = cull height of environmental station antenna.

$JSMS_W$ uses a default value of 30,000 feet (9144 meters) for h_{cull}, which would include most aircraft. Thus, any environmental record more than d_{cull} from the proposed system would be eliminated from further analysis.

The next cull performed is based on a free-space pathloss calculation. A preliminary interference-to-noise ratio (I/N) calculation is made, ignoring frequency-dependent rejection (FDR), and the result is compared to the I/N threshold value. If the result is less than the threshold, the environmental record would be eliminated from further analysis.

Step 5 - Compute Received Interference Power Level

The received interference power I, in dBW, within the victim receiver passband from an interfering source transmitter, is calculated using:

$$I = 10 \log (P_T) + G_T - L_{CT} + G_R - L_{CR} - L_P - L_{POL} - FDR \quad \text{(Equation 2)}$$

where

P_T = interference source transmitter power, in watts

G_T = interference source antenna gain in the direction of the receiving antenna, in dBi

G_R = victim receiver antenna gain in the direction of the interference source antenna, in dBi

L_{CT} = cable/insertion loss of the transmitter system, in dB

L_{CR} = cable/insertion loss of the receiver system, in dB

L_P = propagation path loss, in dB

L_{POL} = antenna polarization mismatch loss, in dB

FDR =frequency-dependent rejection, in dB

A description of the calculation/specification of each term in Equation 2 is given in the paragraphs that follow.

Transmitter Power (P_T)

The transmitter power, P_T, is either accessed directly from the environmental database for existing systems or provided by the user for the proposed system. User-controlled default values for the transmit power, based on frequency band, are used in the event that the data field is empty.

Antenna Coupling ($G_T + G_R$)

The antenna coupling between the interfering transmitter and the victim receiver is computed using:

1. The mainbeam gains and azimuth orientations of the two antennas, which are available from the environmental database or proposal record. Default values for antenna gains are used in the event the data

field is empty.

2. The relative orientation of the transmitter and receiver antennas, which is computed using transmitter and receiver antenna coordinates from the environmental database or proposal record.

3. Statistical antenna-gain algorithms included within the model, which give the antenna gain as a function of off-axis angle for a given mainbeam antenna gain.

Note that if the transmitter system is mobile, the antenna coupling G_T, is the peak gain for that system; similarly, if the receiver system is mobile, the antenna coupling G_R is the peak gain for that system.

Figure 1 is an example of a transmitter/receiver configuration. The antenna azimuth angles, θ_1 and θ_2, are specified with respect to true north. First, the orientation angle, ϕ, of the receiver relative to the transmitter is computed. $JSMS_W$ uses an approximation rather than the rigorous definition for a spherical surface in order to minimize execution time.

Specifically,

$$\Delta_{LAT} = T_{LAT} - R_{LAT} \qquad \text{(Equation 3)}$$

$$\Delta_{LON} = K_1 (T_{LON} - R_{LON}) \qquad \text{(Equation 4)}$$

$$\phi = \arctan(\Delta_{LON}/\Delta_{LAT}) - K_2 (T_{LON} - R_{LON}) \qquad \text{(Equation 5)}$$

where

T_{LAT} = transmitter latitude, in radians
R_{LAT} = receiver latitude, in radians
T_{LON} = transmitter longitude, in radians
R_{LON} = receiver longitude, in radians
K_1 = $[\cos(T_{LAT})\cos(R_{LAT})]^{0.5}$
K_2 = $(T_{LAT} + R_{LAT})/2$
ϕ = orientation angle, in radians

Note that the actual algorithm for calculating ϕ is somewhat more complicated than that in Equation (5) since the actual algorithm must be applicable in all four quadrants and must account for the longitudinal coordinate discontinuity at the international dateline (i.e., the international dateline is +180 degrees longitude when approached from the western hemisphere and -180 degrees longitude when approached from the eastern hemisphere).

This approximation, which includes a correction to account for the fact that the equator is the only closed locus of points at a constant latitude that is a great circle, will provide an azimuth estimation error of less than one degree for systems separated by 300 km in the mid-latitudes.

The distance d, in km, between the two sites is computed using:

$$d = 6378 [\Delta_{LAT}^2 + \Delta_{LON}^2]^{1/2} \qquad \text{(Equation 6)}$$

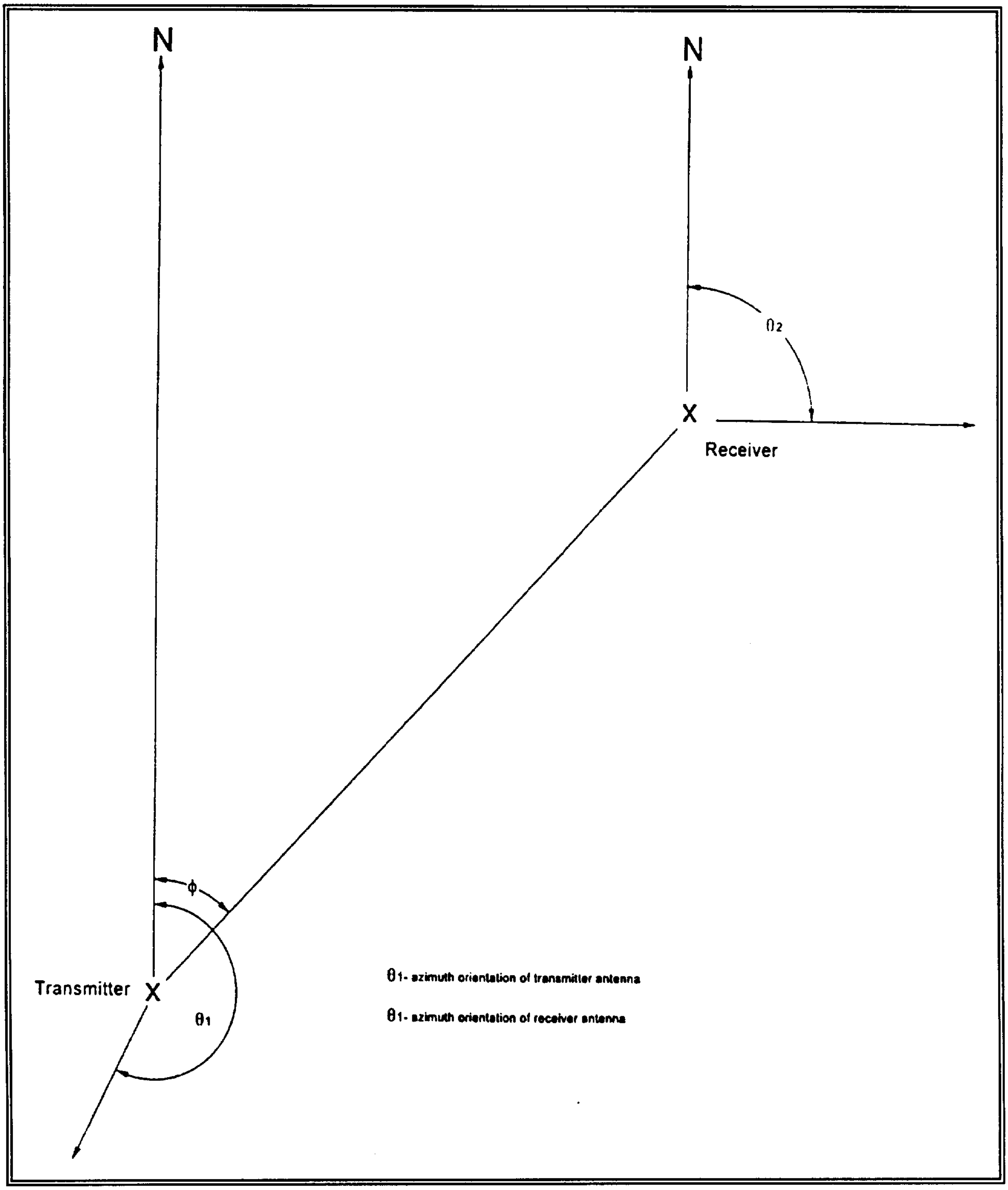


Figure 1. Example Transmitter/Receiver Geographic Configuration.

The off-axis angle to the receiver relative to the mainbeam antenna azimuth angle of the transmitter is given by:

$$\text{off-axis} = \text{Max}(\phi, \theta_1) - \text{Min}(\phi, \theta_1) \qquad \text{(Equation 7)}$$

where the Min and Max operators give the minimum and maximum, respectively, of the elements of the argument. Using this angle and the mainbeam antenna gain of the transmitting antenna, the gain of the transmitter in the direction of the receiver is calculated using one of two statistical algorithms. One is used for antennas with a mainbeam gain ranging from 0 to 9 dBi. Another is used for antennas with mainbeam gains greater than 9 dBi. If the azimuth orientation field is empty, the mainbeam gain is used for all azimuth angles.

This entire process is then repeated after interchanging the coordinates of the transmitter and the receiver to compute the gain of the receiving antenna in the direction of the transmitting antenna. The two off-axis gain values are then summed to get the antenna coupling, $G_T + G_R$.

Polarization Mismatch Loss (L_{POL})

Polarization mismatch loss is only calculated if the victim or interferer is in the mainbeam of the interferer or victim, respectively.

Table 1 provides a listing of the antenna polarization cases implemented within $JSMS_W$, along with their respective designator code. Table 2 gives the polarization mismatch loss that is used for all combinations of polarization.

Table 1. Antenna Polarization Cases and Designators

Antenna Polarization Configuration/Code	Designator
Rotating	D
Elliptical	E
45-degrees	F
Fixed Horizontal	H
Linear	J
Left-hand circular	L
Right-hand circular	R
Horizontal and Vertical	S
Right- and Left-hand Circular	T
Fixed Vertical (Default Value)	V

Table 2. Polarization Mismatch Loss

Transmitter Antenna	Receiver Antenna	Mismatch Loss (dB)
H	V	20
H	F,L,R,T	3
H	E	2
H	J	1
H	S,D,H	0
V	H	20
V	F,L,R,T	3
V	E	2
V	J	1
V	S,D,V	0
L	R	20
L	F,H,J,S,V	3
L	E	1
L	T,D,L	0
R	L	20
R	F,H,J,S,V	1
R	E	1
R	T,D,R	0
E	F,H,J,S,V	2
E	L,R,T	1
E	D,E	0
F	H,L,R,S,T,V	3
F	E	2
F	J	1
F	D,F	0
J	L,R,T	3
J	E	2
J	F,H,S,V	1
J	D,J	0
S	F,L,R,T	3
S	E	2
S	J	1
S	D,H,V,S	0
T	F,H,J,S,V	3
T	E	1
T	D,L,R	0
D	D,E,F,H,J,L,R,S,T,V	0

Cable Insertion Loss (L_{CT}) & (L_{CR})

Both the transmitter and receiver cable insertion losses are set to 2.0 dB.

Propagation Pathloss

In general, the $JSMS_W$ uses TIREM, a spherical earth model (SEM), or free-space propagation to compute the propagation pathloss. TIREM, which is supported by a built-in terrain database, is employed for all path-loss calculations in the 1-MHz to 20-GHz frequency range, provided that terrain data is available. The TIREM model is automatically replaced by the SEM during an analysis if a radius of operation is associated with the transmitter and/or receiver station or if the terrain data needed is unavailable The free-space propagation formula is used outside the 1-MHz to 20-GHz range.

Step 6 - Compute System Noise Power Level

The system noise power includes external environmental noise, transmission line noise, and internal receiver noise. A standard method of computing the total system noise power is to find the equivalent system noise temperature, T_{SYS}, which is equal to the noise temperature of the antenna, T_{ANT}, plus the noise temperature of the receiver, T_R. T_{ANT} accounts for both external environmental and transmission line noise). The total receiver system noise power N, in dBW, is then given by:

$$N = 10 \log (kT_{SYS}B)$$

(Equation 8)

where k = 1.38 x 10^{-23} J/K (Boltzman's constant), $T_{SYS} = T_{ANT} + T_R$, in Kelvin, and B is equal to receiver bandwidth, in Hz. The latter can be effectively represented by the receiver 3-dB bandwidth in most cases, which is set equal to the necessary bandwidth contained in the emission designator since the receiver 3-dB bandwidth information is not available in the current database.

The receiver noise temperature, T_R, is calculated using the standard formula:

$$T_R = T_O (10^{(F/10)} - 1)$$

(Equation 9)

where T_O = 290 K, and F is the receiver noise figure in dB. The latter is automatically assigned based on the emission designator and the frequency of operation.

The antenna noise temperature, T_{ANT}, is given by:

$$T_{ANT} = [290(L_{CR} - 1) + T_{EVMT}]/L_{CR}$$

(Equation 10)

where L_{CR} is equal to the receiver cable loss factor (dimensionless), and T_{EVMT} is the external environmental noise temperature. The latter is determined using

$$T_{EVMT} = \{10^{(N_{EVMT} - 204)/10}\} / k$$

(Equation 11)

where N_{EVMT} is the frequency-dependent environmental noise level, which can be represented as a set of piece-wise continuous functions as shown in Figure 2. The curves, derived from ITU-R study group recommendations, represent measured data for the worst times/worst locations (Noisy/Urban), best times/best locations (Quiet/Rural), and a median range (Average/Suburban). These composite curves account for atmospheric (lightning), man-made (urban, suburban, and rural), and galactic noise sources. The user may select which curve is used by $JSMS_W$, with the default being Median.

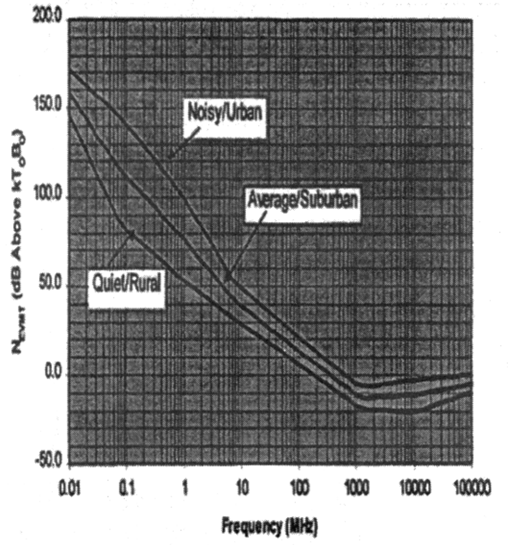


Figure 2. External Environmental Noise Level, N_{EVMT}, as a Function of Frequency

Step 7 - Determine Predicted Interference-Free Frequencies

At this point in the analysis, all terms in Equation 2 have been computed except the FDR. But prior to calculating FDR, the I/N must be compared to the I/N threshold value. If the I/N is less than the threshold, the environment record would be eliminated from further analysis and hence discarded as a possible source, or victim, of interference.

Frequency-Dependent Rejection (FDR)

The FDR accounts for the potential interference power reduction due to the filtering effect of the victim receiver selectivity on the interfering transmitter emissions (e.g., adjacent-channel or wideband-cochannel interference).

The FDR is computed by piece-wise integration of the product of the transmitter emission spectrum and the receiver selectivity.

Transmitter Emission Spectra Representation

Transmitter emission spectra are represented by the emission masks specified in Chapter 5 of the *NTIA Manual* (Reference 1). This representation renders a conservative interference analysis since the actual emissions are contained within the mask to meet spectrum standards.

Station class, transmitter frequency, emission designator (bandwidth and modulation), and transmitter power data from the environmental database or the proposal record are used to automatically select the appropriate emission mask. In the event that the emission designator data field is empty, default values are used. The mask selection algorithm was derived from Chapters 5 and 6 of the *NTIA Manual*.

Receiver Selectivity Representation

Receiver selectivity is modeled as a bandpass filter with the unity gain pass band set equal to the necessary bandwidth contained in the emission designator of the station record, which is typically the 20 to 30 dB emission bandwidth. This is somewhat conservative, but there is no other bandwidth data available in the environmental database. The selectivity roll-off slope employed is based on the emission designator and frequency of operation. Default slope values are automatically selected for different emission types and frequency ranges.

Calculation of FDR

The received power is computed by piecewise integration of the product of the power spectral density (PSD) at the receiver input and receiver selectivity. Specifically, the product is partitioned into regions where each of the two curves has a constant slope in dB/Hz for that section. The equations for both line segments are then added; this is equivalent to taking the product of the two functions when they are expressed in conventional units. The closed-form integral of the sum is then used to give the received power for that segment. The integral is easily evaluated since the roll-off is modeled as linear-in-frequency. When viewed on semi-log paper, the linear-in-frequency roll-off is convex about a logarithmic roll-off curve. The received power is thus greater for a linear roll-off as compared to a logarithmic roll-off.

The FDR algorithm is used in an iterative process in order to determine the frequency offset (delta-F) required to ensure an interference-free operation. As each new FDR value is computed, the I/N is calculated and subsequently compared to the I/N threshold value. As soon as the threshold has been met, the process stops and the delta-F value is then used as a guardband around the environmental record's frequency. Thus, use of some of the subband of the candidate frequency range is denied due to the potential interference related to the environmental record being analyzed.

Step 8 - Rank the Interference-Free Frequencies

For each frequency that meets the criteria, i.e., tuning increment, allocation or allotment plan constraints, and the I/N threshold, a rank is determined. The ranking criterion is that the most constrained frequency should be the first one recommended for assignment. The acceptable frequencies are ranked initially by reuse, i.e., the frequency with the most assignments is ranked first. If the reuse values are equal, the edge number is used. Each nominated frequency is given an edge number of zero, one, or two that indicates the constraints relative to using adjacent frequencies as shown in Figure 3, where channels 3, 8, and 11 represent nominated frequencies with edge numbers of 0, 1, and 2, respectively. Those frequencies with equal reuse numbers are then ranked in descending order of edge number in order to rank frequencies near the edges of occupied spectrum, higher than those frequencies in the center of assignment-free spectrum.

When each of the eight steps have been completed, the frequency assigner can assign frequencies to proposals and then submit the proposals for the appropriate national-approval process. A nomination analysis results screen is shown as Figure 4.

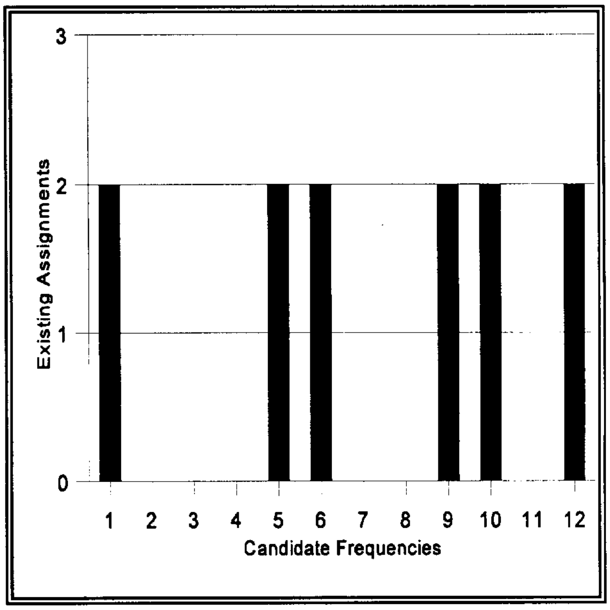


Figure 3. Candidate Frequencies with Count of Existing Assignments

INTERFERENCE ANALYSIS OUTPUT REPORT
ENVIRONMENT DATA

Frequency : M30.9 to M31.2 | J/12 Equipment :
TX Location : AUSTIN | Emission : 16K00F3E
TX Coordinates : 300333N 0974842W | Tuning Increment : K25
TX Auth Radius : 0 | Transmitter Power : W35
RX Location : AUSTIN | TX Antenna Gain : 3.00
RX Coordinates : 300333N 0974842W | Station Class : FBR

Type of Analysis : IA Level (TIREM/SEM)
Display SPACE/AREA Flags : YES
Environment I/N Threshold (dB) : 0.00
Proposal I/N Threshold (dB) : 0.00

Flags: snotes, space, area, coordinates

Rank	Frequency	Service	Flags
1	M31	ALLOCATION TABLES NOT CONSIDERED	----
2	M30.95	ALLOCATION TABLES NOT CONSIDERED	----
3	M30.925	ALLOCATION TABLES NOT CONSIDERED	----
4	M31.075	ALLOCATION TABLES NOT CONSIDERED	----
5	M31.125	ALLOCATION TABLES NOT CONSIDERED	----
6	M30.975	ALLOCATION TABLES NOT CONSIDERED	----
7	M31.1	ALLOCATION TABLES NOT CONSIDERED	----

Figure 4. Nomination Results

REFERENCES

[1]U.S. Department of Commerce, NTIA, *Manual of Regulations and Procedures for Federal Radio Frequency Management*, Washington, DC, September 1995 with Revisions for September 1996 and January and May 1997.

[2]Keech, T., O'Hehir, M., and Hensler, T., *JSMS$_W$ Interference Analysis Algorithms*, JSC-CR-96-016B, DoD JSC, Annapolis, MD, April 1998.

PAPER TITLE : **Paper 3 - Spectrum Management Using JSMSw**

AUTHOR : T. Hensler

NAME : K.S. Kho

QUESTION :

How flexible is the assignment regime? Is it capable to assign all systems working in the band you mentioned. Could it assign for instance A/G/A frequencies?

ANSWER :

JSMSw is flexible, however, it is best for ground-to-ground assignments. A possible future enhancement would be to permit users to select the type of intersite analysis to be performed. JSMSw is modular and it would be possible to tailor the analysis to the frequency band and user.

PAPER TITLE : **Paper 3 - Spectrum Management Using JSMSw**

AUTHOR : T. Hensler

NAME : M. Elliott

QUESTION :

In order to perform interference analysis on a frequency proposal calculation, it is obviously necessary to have a background database of existing assignments. Could you please describe how such data can be entered into the JSMSw tool?

ANSWER :

All data entered in JSMSw is via vertical SFAF. We have a conversion programme for GMF data and could write conversion routines for other data formats, e.g. NATO 14. Point. The JSC provides CD ROMs containing FRRS, FCC and ITU data with routines for record selection into SFAF files for import into JSMSw.

NAME : D. Jaeger

QUESTION :

Is there any reason for limitation of the programme just at 1GHz? Many problems in the aircraft business will start in the frequency band just above.

ANSWER :

There is no limit at 1 GHz. In the US, the most used frequencies are between 30 MHz and 1GHz. The algorithm is valid for frequencies from 2 MHz to 20 GHz.

HF FREQUENCIES - SHARING AMONG NATO COUNTRIES

M. Proia
G. Maviglia
ELMER S.p.A
Viale dell'Industria 4
00040 Pomezia (Rome) Italy

SUMMARY

In this paper, an efficient use of HF frequencies (2MHz to 30MHz) is investigated to establish radiocommunication between moving platforms operating in a given area and ground-based fixed stations suitably located to provide effective coverage of the area.

In the first part of this paper, the use of the HF spectrum (2MHz to 30MHz) as a primary communication resource is justified, taking into consideration that the constraints associated with ionospheric propagation and spectrum congestion require efficient use of the available frequencies.

These constraints may produce degradation in the skyway links.

Considering the problem of communication between mobile units and ground-fixed stations in a NATO environment, the following part of the paper contains a description of the basic concepts followed by the ELMER engineering activity in the development of an HF radio communication system intended to allow any mobile unit to establish a link with at least one ground station within a predefined pool.
This is achieved through mapping of the communication area by ground stations suitably located, combined with an efficient frequency management.

An application is described in which the system concept is validated of the Mediterranean area.

It should be understood, however, that the results of the investigation can be validated for any other area, with some modifications, if any.

LIST OF SYMBOLS

HF	High Frequency
NATO	North Atlantic Treaty Organisation
BLOS	Beyond Line Of Sight
UHF	Ultra High Frequency
SHF	Super High Frequency
ALE	Automatic Link Establishment
RF	Radio Frequency
LUF	Lowest Usable Frequency
MUF	Maximum Usable Frequency
OWF	Optimum Working Frequency
ASAPS	Advanced Stand Alone Prediction System
GPS	Global Positioning System
TDMA	Time Division Multiple Access

1. HF VS SATELLITE COMMUNICATION

Until the advent of satellite communication, the HF channel has practically been the only resource available to establish point-to-point links of the BLOS type (Beyond Line Of Sight) and it is still extensively used as a primary means of communication to meet the requirements of military and civil Agencies of many Countries all over the world.

In line with the main area of interest and operational experience of the ELMER Company, this paper is focused on military communication requirements, even though the concepts from which the system architecture is derived are equally applicable to civil applications.
In order to put into the correct perspective the use of the HF channel as a medium/long range communication support, a brief comparison is given below between the HF and the satellite channels operating in the UHF and SHF bands.

Both channels are capable of supporting long-distance communication between moving platforms and ground stations in a variety of services ranging from repetitive broadcasts of simple texts on teletype to secure digital voice/data .

As a general consideration, satellites are vulnerable and this is a major issue in maintaining HF as a primary resource rather than a backup for military applications.
Furthermore, the implementation of satellite communication facilities onboard moving platforms (aircraft in particular), poses a number of technical and economical problems, such as :

- Acquisition of new, dedicated radio equipment to be added to the communication resources already existing onboard the platform.
This has an economical impact common to all the installations and a technical impact for the small platforms, in particular light mobile unit, in which additional weight and installation complexity are to be avoided as much as possible.

- Use of high-gain, steerable antennas for satellite tracking. Again, this problem assumes particular relevance for high-dynamics platforms, in particular aircraft.

- Overhead costs associated with the use of a satellite channel.
These disadvantages are compensated by the fact that satellite communication has a high level of reliability, since the links are "line-of-sight" and largely immune to changes in the physical parameters of the propagation channel.

Paper presented at the RTO IST Symposium on "Frequency Assignment, Sharing and Conservation in Systems (Aerospace)", held in Aalborg, Denmark, 5-7 October 1998, and published in RTO MP-13.

On the other hand, use of the HF band for long-range communication is the traditional approach which combines cost effectiveness with ease of implementation as a result of :

- free access to the communication channel
- use of radio equipment currently available onboard aircraft and ship
- simple installation
- low procurement cost when compared with other types of long-range communication systems.

These advantages are accompanied by some traditional constraints related to ionospheric propagation, spectrum congestion, the intrinsic low capacity of the HF channel when compared with other communication channels and the need of user expertise in the operation of manually controlled systems.

In these recent years, however, a number of significant technological advancements have been implemented in the HF communication systems, among them the ALE (Automatic Link Establishment) procedures, which automatically select the best frequency for the channel conditions.

In many countries , HF systems are still the backbone of long/medium range military and civil communications , primarily to support ship-shore ship and ground-air-ground links.

HF systems are also used within public telephone networks to extend coverage to remote areas in which use of the VHF and UHF frequencies is precluded by unfavourable siting conditions.

2. FREQUENCY PREDICTIONS

Long-distance communications over the HF channel take avail of RF frequency predictions derived from probing of the ionosphere parameters and modelling of the complex phenomena involved in ionospheric propagation.

Data on the physical conditions of the ionosphere are obtained by probing with RF pulses transmitted and received by ground stations or direct measurements made on the ionosphere layers by instrumentation installed onboard rockets.
RF frequency predictions based on probing and modelling of the ionosphere are available from scientific institutions all over the world.
These predictions are currently used in the preparation of frequency plans and in the assignment of frequencies intended to support communication over a given area.

The predictions indicate that for each link there is a Maximum Usable Frequency (MUF) and a Lowest Usable Frequency (LUF) which define the band of usable frequencies. In this band, an Optimum Working Frequency (OWF) is also identified. Typically, OWF = 0.9 MUF.

The predicted values of the MUF, LUF and OWF frequencies change in a "regular" way due to the "regular" variations of the ionospheric parameters over the day, the year and the 11-year sunspot cycle.

Consequently, the communicator will select, among the assigned pool, one frequency that is expected to be suitable to support the desired link at the given time.

However, unpredictable changes may occur, caused by unpredictable perturbations of the ionosphere, which could preclude use of the frequency selected for the link.

3. THE HF CHANNEL: THE PROBLEMS AND THE POTENTIAL SOLUTION

As previously anticipated, in the use of HF frequencies (2MHz to 30MHz) for radio communication over long/medium distance links, the system designer must face problems related to the characteristics of ionospheric propagation and to spectrum congestion.

The behaviour of the HF channel as a communication resource can be summarised in the following points:

a) The practically unlimited access to the HF band by a large number of users and the long-range propagation characteristics of the HF frequencies create spectrum congestion, which reduces the number of usable frequencies.

b) Due to the physical phenomena associated with skywave propagation over the HF channel, the frequency capable of supporting a given link depends on the link distance and on the seasonal/ monthly /daily variations of the propagation parameters.
More specifically, assuming single-hop propagation, one frequency is capable of supporting communication only over a certain distance range or area at a certain time.

c) The establishment of a link depends both on the conditions of the ionosphere and the operative parameters of the system; in particular, the RF power of the transmitter and the efficiency of the radiation system. This last parameter depends on the ratio of the physical dimensions of the radiation system to the wavelength (of the RF signal).

d) After a link has been established, sudden changes may occur in the parameters of HF propagation, mostly at dawn and sunset, such to produce an intolerable degradation and possibly complete loss of communication.

In this situation, link re-establishment is attempted by switching to another frequency, if available.

The choice of the correct frequency capable of supporting a given link depends on the distance between the two correspondents, their geographical locations, and the time-variable conditions of the ionospheric layers.

Assuming single-hop conditions, ionospheric propagation shows a selective behaviour in frequency for a fixed distance

and vice versa a selective behaviour in distance for a fixed frequency.
This implies that a specific frequency (or more generally a restricted set of frequencies) be used to support communication over a given distance at a given time and, conversely, only a certain distance can be covered by using a specific frequency.

This concept is illustrated in the following Figures 1.a and 1.b.

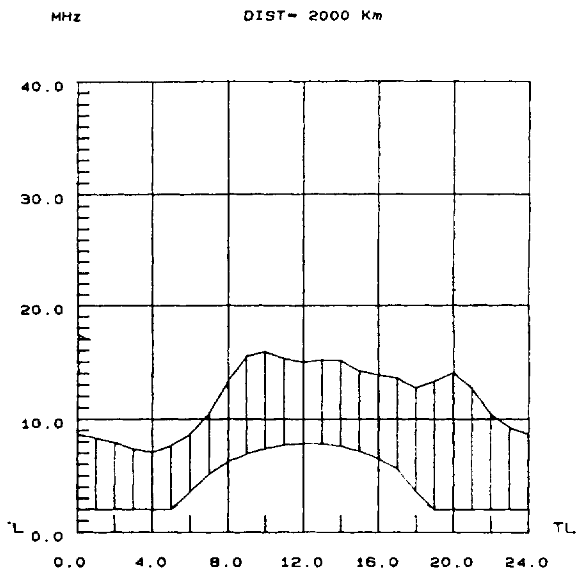


Figure 1.a: This Figure shows the selective behaviour in frequency for a fixed distance (2000 Km).

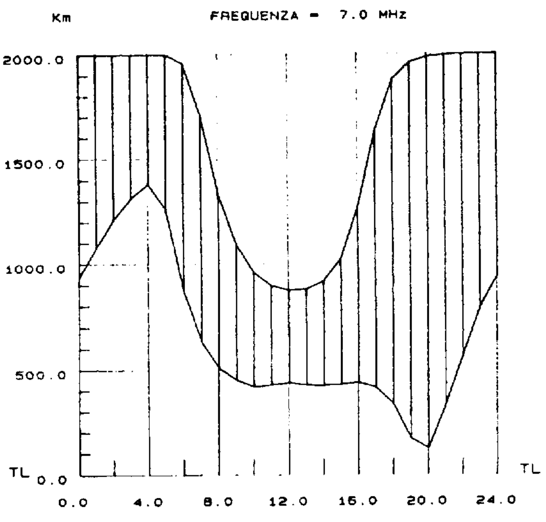


Figure 1.b: This Figure shows the selective behaviour in distance for a fixed frequency (7.0 MHz).

Even though the correct frequency is used, unpredictable changes may occur in the parameters of HF propagation such to produce an intolerable degradation and possibly complete loss of communication with the desired correspondent.

In a scenario of communication between mobile units and ground stations, the mobile unit may not be capable of establishing a link with the designated station. The approach followed in the system described in this paper is to allow the mobile unit to establish a link with other ground stations, acting as intermediaries, if the directly link is not available.

This concept finds an easy implementation in the NATO environment where the ground stations are interconnected over existing infrastructures, partly owned by NATO and partly by the Nations, using different transmission supports (wire, microwave, satellite, etc.).

In order to enable a mobile unit to link with different ground stations, it is necessary to assign to each ground station a coverage area with the following criteria:

- The coverage areas of the individual ground stations must extended as much as possible over the entire NATO territory.
- Overlap of the coverage areas should be maximised.

4. THE SCENARIO

HF communications between a mobile unit and ground stations over a given area requires that a number of predicted frequencies be available, capable of supporting links over different distances at different times over the 24 hours.

It is clear that to increase communication reliability, the frequency plan should include the largest number of predicted frequencies. The ALE (Automatic Link Establishment) function is a typical example of use of a large number of frequencies among which the one suitable to support the desired link is selected.
In principle, the ALE technique automatically scans a set of frequencies to produce a priority list of frequencies based on the quality of the exchanged signal.

However the HF frequency band is a limited resource and the frequency planning activity must therefore face the problem of efficient spectrum utilisation, by implementing a frequencies sharing scheme.
As previously anticipated, the NATO area could be covered by a number of different ground stations on the NATO territory; a good coverage of the complete area and a good overlap, can be obtained by using a suitable number of ground stations. Each ground stations is associated with a set of frequencies which allows connection in the coverage area of that ground stations.

In this condition any mobile unit shall be enable to use all the sets of frequencies of the ground stations as necessary, depending on the sub-area in which the mobile unit is located.

More specifically the coverage zone of a ground stations is defined by its location in the NATO area and its extension.

The location is the result of a mapping which assigns to each station a coverage distance. The extension of the coverage zone depends on the operative parameters of the link (signal to noise ratio, probability of success of link, etc.).

In this situation any mobile unit in the NATO territory is enabled to establish a reliable link with at least one ground stations.
The objective of the system designer is to obtain good coverage overlap within the constrains of efficient use of the spectrum.

In order to achieve the best compromise is necessary to conduct an investigation on the locations of the ground stations in the NATO area (existing and future) and a study to optimise the assignment of coverage area to such ground stations.

As it concerns system operation, each mobile unit shall be assigned a set of frequencies capable of establish link over the NATO area. In these conditions the mobile unit will select the frequency most suitable to support a link with the available ground stations.

In this scenario the mobile unit, if uncapable of linking with desired station, shall be able to establish connection with another ground stations which will handle the traffic between the mobile unit and final destination using the existing infrastructures.

As an application, a study has been conducted by the ELMER system engineering activity on the implementation of the described system concept in the Mediterranean area, with the use of experimental data on ionospheric propagation in support of the concept feasibility.

The system concept can be validated for the entire NATO area by conducting studies and experimentation aimed primarily at achieving data on the physical behaviour of the ionosphere in non-Mediterranean areas.

5. THE APPLICATION

In this paragraph the application is described by indicating the principal results.

The two presentations in Figure 2 are HF frequency predictions for a fixed set of values of the link parameters (distance, solar activity, transmitter power, signal to noise ratio at the receiver).
Figure 2.a shows a map of the Mediterranean area with a fixed station located in Rome and a set of curves that show the skip-distances for the indicated link parameters. The skip-distance, relative to a given frequency, is the minimum distance at which is possible to establish a radio communication at that frequency over the ionospheric channel.
It is important to note that the skip-distance curves for the different frequencies are approximately circular with centre in the transmitter station.

In Figure 2.b is defined a zone that represents a geographic area covered with ionospheric propagation by a fixed transmitter terminal working at fixed frequency of 9MHz. The width of the area is defined by the following parameters: a preset value of RF power, minimum receiver field, the antenna pattern, the time of day, for fixed ionospheric conditions.

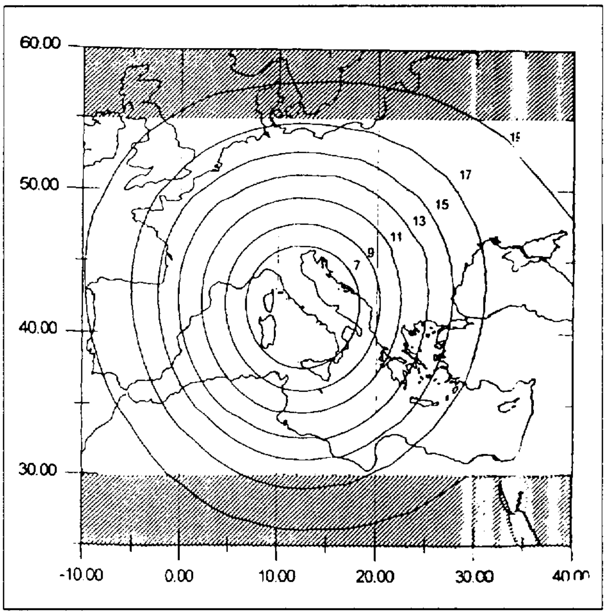


Figure 2.a: Communication distances and associated frequencies showing skip distances for different frequencies, for a preset value of RF power at a give time of day.

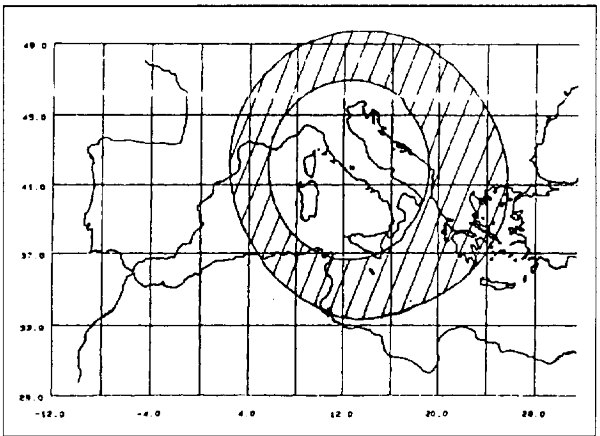


Figure 2.b: Coverage area for a fixed frequency (9MHz) with a preset value of RF power at a given time of day.

The approximately circular shape of the curves in Figures 2.a and 2.b seems to indicate that the composition of the ionosphere over the Mediterranean area is quite homogeneous.

The effect of the ionospheric layer on the impinging radio wave can be visualised as a reflection in accordance with the law of optics that the angle of arrival is equal to the take-off angle. This implies that the same frequency can be used at a given time to support links over the same distance. irrespective of the geographical coordinates of the two correspondents.

This very important result can be validated by use of the software simulation program (ASAPS), which indicates that, for the same values of distance, solar activity, and transmitter power, the same signal to noise ratio is obtained at the receiver for the same link frequency.

It would be interest to know if the homogeneous composition of the ionosphere extends beyond the limits over the Mediterranean area. This requires that investigations be conducted possibly through NATO teamwork.

At this point, the problem is to obtain a coverage of the desired area with a minimum number of ground stations, each provided with a minimum number of frequencies.

As told before, the trade-off is between efficient coverage and spectrum conservation, associated with the cost for the implementation of the ground stations.

The frequency minimisation scheme adopted in the HF system described in this paper is based on ionospheric observations made by the National Institute of Geophysics in Rome (Italy) and a software simulation program (ASAPS). The observations also are important to establish the extension of the area covered by a fixed station (see Figure 2.b) and to demonstrate that a fixed frequency is capable of supporting radio communication between the ground station and mobile units located on the nearest and most distant points of the ring-shaped zone, with a degradation of the signal to noise ratio at the receiver not to exceed a fixed value below the value at the middle of the zone, as indicated in Figure 3.

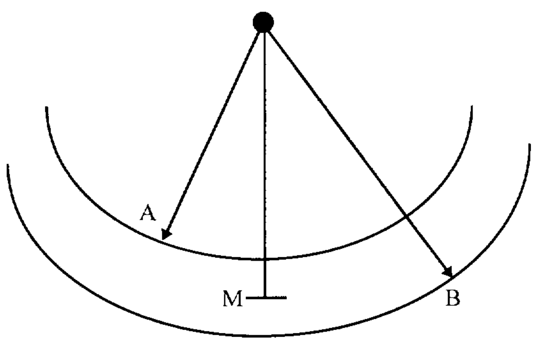


Figure 3: Representation of a ring-shaped zone and ground station. The width of the zone is established by the acceptable degradation of signal to noise ratio at the nearest point (A) and most distant point (B), with respect to the value at the middle point (M).

The observations are based on a channel probing experimentation conducted to define the coverage of the Mediterranean area provided by a ground-based fixed stations, under conditions of single-hop ionospheric links via E layer or F layer, solar activity: sun spot number of 132 (0÷250), transmitter RF power of 400 Watt, signal to noise ratio at the receiver ≥10dB and probability of connection >90%. The fixed value of acceptable range of degradation of signal to noise ratio is 2dB.

A selected set of data derived from ionospheric predictions is provided in attachment A.
It is of interest to note that for the same distance, the ionospheric and operative characteristics are practically the same, regardless of the coordinates of the correspondents.

On the basis of these considerations, a computer analysis on HF propagation in the Mediterranean area has been conducted using the ionospheric prediction programs in order to obtain frequencies usable for point-to-point links.

The first step is to set up a system configuration that will allow a coverage of the Mediterranean area as uniform as possible with a minimum number of ground stations.

The Mediterranean area is defined as the area extending over the 10W to 40E meridians and 30N to 50N parallels.
The results of the observations made by the ASAPS software program and National Institute of Geophysics in Rome, suggests a mapping of the coverage area into ring-shaped zones centred in the location of each ground-station.

To obtain a good mapping of the Mediterranean area through sub-areas assigned to the different ground-based stations, it is reasonable to consider distances 500km and/or 750km and/or 1000km from the stations.
The goal is to map the area so that it can be possible from any point to establish a link with one (and possibly more) ground-station by using the assigned frequencies.

Considering the geographical position of the Italian peninsula in the Mediterranean, it seems appropriate to think of a mapping consisting of annular zones extending into the Mediterranean area, centred around ground stations already located in appropriate sites along Italy's coastline.

The concept is to allow any mobile units to establish a link with a ground station located at a distance requiring a frequency which is the best for ionospheric propagation at that time; for example, to establish a reliable link at dawn (i.e. at a time when the good frequencies are in the lowest part of the HF spectrum) a station located at a longer distance must be selected as correspondent.

This point needs further elaboration. Under fixed propagation conditions at a given time, it may happen that the interval of frequencies usable to establish a link at a given distance is narrow and not adequate to support even small variations of the ionosphere parameters. This situation is often encountered at dawn and sunset, as indicated in this Figure 4.a.

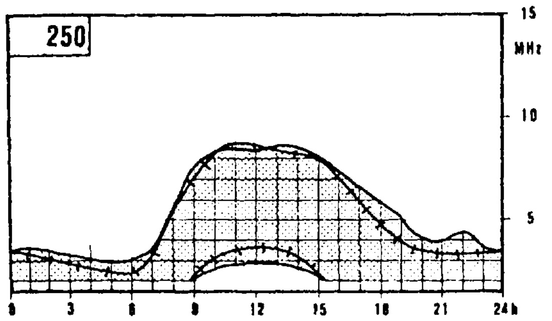


Figure 4.a: The Figure shows the MUF/LUF diagram for a fixed propagation conditions and for a fixed distance (250 Km).

To increase the width of the frequency interval, two alternatives are available:
* increase the RF transmitter power for fixed distance
* increase the link distance for fixed power

The increase of RF power is not feasible in most cases and consequently the second alternative is to be followed.
The advantages related to the increase of the link distance are visualised in Figure 4.b.

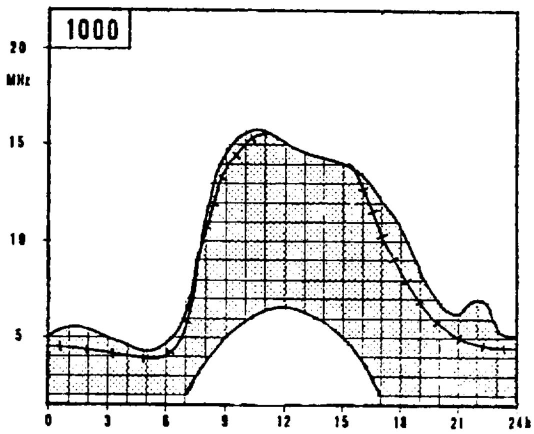


Figure 4.b: The Figure shows the Muf/Luf diagram for the same propagation conditions of Figure 4.a but for different distance (1000 Km). The MUF/LUF diagram is more large than diagram of Figure 4.a in particular at dawn and sunset.

In Figure 5 this concept is visualised in a qualitative manner by showing the coverage of a annular zone at a distance of 500km from a station located on the Adriatic coast of Central Italy. By use of a map of Italy, the location of this station can be identified in the vicinity of Ancona.

As a result of computer modelling, a network configuration consisting of four stations to be located at Trieste, Ancona, Rome and Capo Teulada (Sardinia) has been defined as adequate to provide a good coverage of the desired area.

This configuration of ground-stations shown in Figure 6 is an example, specific to the requirements of the Italian Navy; the positions of the stations may change in accordance with different system and area coverage requirements.

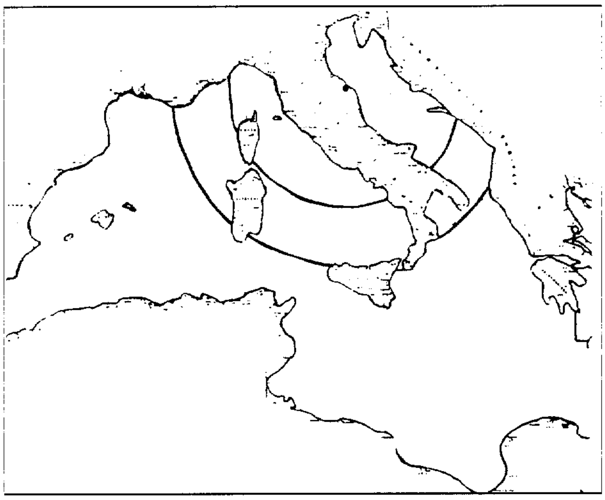


Figure 5: area covered from the ground-station Ancona with distance range of 500km.

It should be observed that in the coverage area of the configuration of Figure 6 there are zones in which a mobile unit is able to establish a link with more than one station. This is possible because of a multiple coverage of such zones; in which it is possible to use more than one frequency to establish links with different stations.

The choice of the frequency to be used among those available depends on a number of different factors, among others :
- the location of the ground station with which the mobile unit must establish a link
- which of the ground stations is not already engaged in a link and is therefore available for communication.

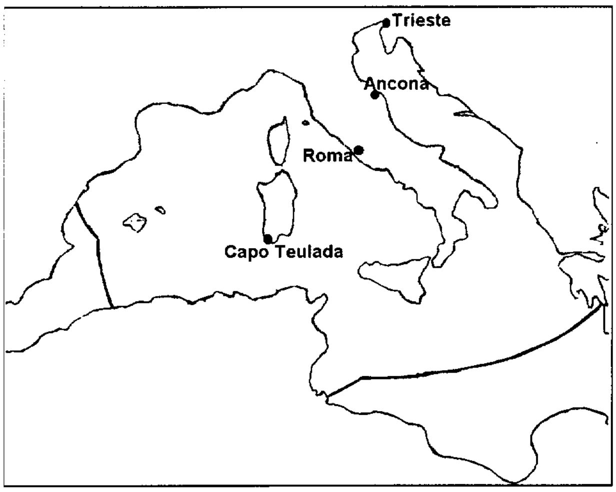


Figure 6: area covered from the ground-stations Ancona, Rome, Trieste and Capo Teulada with distance range respectively of 500km, 500km, 1000km and 500km.

After the minimum number of ground stations of the system has been established, it is necessary to evaluate the operative characteristics of the system.

First of all, it is necessary to perform a prediction of the optimum frequencies to be included in the frequency plan, which will be assigned to the ground stations and mobile units.

In order to know its distance from the ground stations, each mobile unit must know its position and this can be achieved by use of a GPS (Global Positioning System) receiver; on the basis of the GPS information and the list of the assigned frequencies, the correspondent station and the frequency capable of supporting the link will be identified.

The frequencies are selected on the basis of data provided by the HF propagation prediction programs currently available. At hardware level an advantage of the proposed system is that the system function can be integrated in a external Applique interfaced with conventional HF equipment for both ground stations and mobile units.

This allows adoption of the system by all NATO countries with system's costs limited to the procurement of the Applique.

In the following, the major operating functions of the Applique are briefly described.

Connectivity between the mobile units and a subscriber of a ground-based communication system can be established by connecting the ground stations to a computerised Master Station which controls the traffic exchange over the existing infrastructures.
This communication support is necessary because a mobile unit must be provided with capabilities to establish a link with any ground station .
Via the infrastructure net it is possible to reach the final destination.
The intermediaries only need to know the routing data contained in the overhead of the message in plain format, while the information contents of the message can be protected by encryption.

A feasibility study has been conducted on the use of this approach in the implementation of an HF ground-air-ground communication system over the Mediterranean area, for exchange of plain/encrypted voice traffic and mobile unit status/position monitoring messages in a digital format.

As a result of computer modelling, a network configuration consisting of four stations to be located at Trieste, Ancona, Rome and Capo Teulada has been defined as adequate to provide a good coverage of the desired area.
The Master Station provides :

- interface to the Public Telephone Network for message routing
- assignment of the frequencies to the four stations and to the mobile units

The frequencies are computed on the basis of data provided by the HF propagation prediction programs currently available.

The status/position monitoring messages are automatically generated by the mobile unit and include information provided by the mobile unit sensors. The mobile unit access to the channel for transmission of the monitoring messages is by TDMA (Time Division Multiple Access) oriented techniques.
The 'time plan', which assigns to each mobile unit the time slots in which to establish a link, is an output of a prototype software programme currently available; the inputs are the number of mobile units, the number of ground stations, the contents and length of messages, the characteristics of the mobile units (fast or slow in connection with speed), and the operative characteristics of the link (type of modulation, type of coding, baud rate, etc.).

This status/positioning monitoring message is processed by a dedicated receiver provided in each ground station. This receiver is also used to receive any distress messages originated by the mobile units and for this important reason operation of the receiver is assigned for maximum 10% of the time (6 minutes per hour) for receive the monitoring messages. The configuration of the ground station includes:
* one RX/TX for traffic
* one RX for monitoring and distress messages.
As it concerns the mobile unit, a single transceiver is used for booth traffic and monitoring/distress messages.

In addition to the system aspects, the study has also considered the impact of this concept on the standard HF communication equipment, both ground-based and installed onboard the mobile unit.

The results indicate that all the additional functions required can be integrated into an external Applique to the radio equipment.
Consequently, the modifications required in these equipment only consist in the addition of a suitable interface for the external Applique.

The feasibility study of an Applique for use in association with ELMER HF Transceivers has also been included in the study.
In Figure 7 is indicated the interface of an Applique for mobile unit and in Figure 8 the interface of an Applique for ground stations.

A few words to describe the functions of the Applique units. The mobile Applique shall calculate the correct frequency for the link on the basis of the distance between the mobile unit and the ground stations, determined by using GPS information.

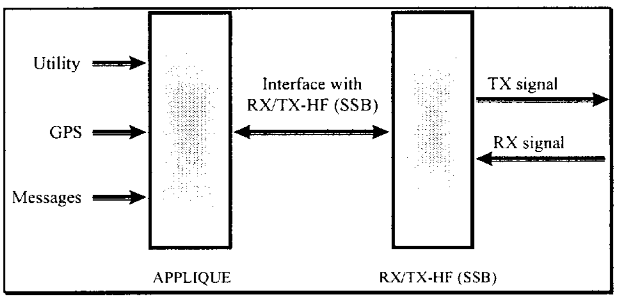


Figure 7: Block diagram of 'mobile system'

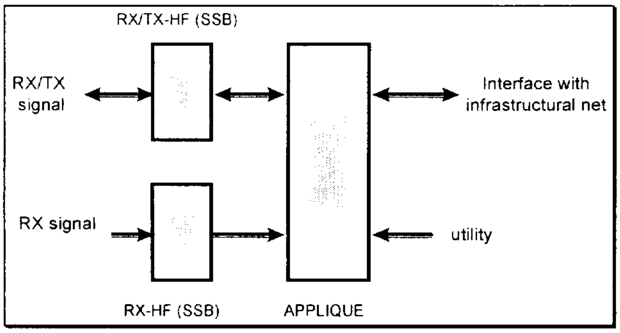


Figure 8: Block diagram of 'ground system'

In particular, the mobile Applique produces a message which contains some information on mobile units about its status/position.

The ground Applique is provided with interface with the associated RF Transceiver and with the infrastructural net.

In conclusion, the approach described in this paper opens the way to new investigations on an optimised use of HF frequencies to ensure reliable long-range communication between mobile platforms and ground-based stations over a NATO countries.

The results of the feasibility study indicate that the HF propagation data provided by the prediction programs currently available allow the HF system designer to achieve an optimised system configuration in which the number of ground stations necessary to achieve the desired coverage can be minimised and spectrum conservation is made possible by a good sharing of the available frequencies.

6. REFERENCES

[01] Kenneth Davies (1969), *"Ionosphere Radio Waves"*, Blaisdell Publishing Co. Waltham, Toronto.

[02] J.M.Kelso (1964), *"Radio Ray Propagation in the Ionosphere"*, Mc Graw Hill Book Comp., NY.

[03] Prediction and Retrospective ionospheric modelling (1992), *"Data validation of ionospheric modelling"*.

[04] International Conference on HF Radio Systems and Techniques (1994), *"HF radio systems and techniques"*

[05] Kelley, Michael C. (1989), *"The earth's ionosphere; plasma physics"*

[06] Gurevich, Aleksandr Viktorovich (1978), *"Non linear phenomena in the ionosphere"*.

[07] Giuseppe Cucinotta (1981), *"Sistemi di comunicazione e interferenze"*, ST.E.T.I S.p.A.

[08] F. Valdoni (1996), *"Appunti di Comunicazioni Elettriche"*, dispense Ing. Elettronica, II Università di Roma "Tor Vergata".

[09] F. Vatalaro (1996), *"Appunti di Radiotecnica"*, dispense Ing. Elettronica, II Università di Roma "Tor Vergata".

[10] Herbert TAUB, Donald L. Schilling, *"Principles of Communication Systems"*, McGraw-Hill Publishing Company, (Second edition).

7. ATTACHMENT A.

ASAPS V2.23 Field Strength Prediction -------------------------

Circuit 1: P2 Distance: 622 KM Date: 15 Aug 1991
Tx: P2 41.45 19.95 Bearings: 277 92 T-index: 125
Rx: Roma 41.90 12.48 Path: Short Path TxAntenna: WH25
FSet: HF-ES2 Noise: -148 dBW/Hz RxAntenna: ISOTROPIC
Mode: 1E Power: 400W BandWidth: 3.0kHz
SIGNAL/NOISE

UT	2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0	11.0
00	..	..	..	..	..	..	..	..	..	..
01	..	..	..	..	..	..	..	..	..	..
02	..	..	..	..	..	..	..	..	..	..
03	..	..	..	..	..	..	..	..	..	..
04	..	..	..	..	..	..	..	..	..	..
05	..	..	13	15	16	..	..	..	..	..
06	..	..	8	12	14	15	..	..	..	..
07	..	..	..	8	11	14	15		..	..
08	..	..	..	9	13	16	18	20	..	..
09	..	..	..	8	12	16	19	18	..	..
10	..	..	..	6	11	15	18	19	21	..
11	..	..	..	6	10	15	17	19	21	..
12	..	..	..	6	10	14	16	18	20	..
13	..	..	..	5	9	13	15	15	..	..
14	..	..	..	8	11	15	16	18	..	..
15	..	..	..	11	14	17	18	..	..	..
16	..	..	7	11	13	15	..	..	..	..
17	..	..	5	9	..	..	..	..	..	..
18	..	..	..	..	..	..	..	..	..	..
19	..	..	..	..	..	..	..	..	..	..
20	..	..	..	..	..	..	..	..	..	..
21	..	..	..	..	..	..	..	..	..	..
22	..	..	..	..	..	..	..	..	..	..
23	..	..	..	..	..	..	..	..	..	..

Table 1: This table shows the S/N value for a link whose parameters are indicated in the heading of this table.

ASAPS V2.23 Field Strength Prediction -------------------------

Circuit 1: P3 Distance: 623 KM Date: 15 Aug 1991
Tx: P3 37.90 17.60 Bearings: 317 133 T-index: 125
Rx: Roma 41.90 12.48 Path: Short Path TxAntenna: WH25
FSet: HF-ES2 Noise: -148 dBW/Hz RxAntenna: ISOTROPIC
Mode: 1E Power: 400W BandWidth: 3.0kHz
SIGNAL/NOISE

UT	2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0	11.0
00	..	..	..	..	..	..	..	..	..	..
01	..	..	..	..	..	..	..	..	..	..
02	..	..	..	..	..	..	..	..	..	..
03	..	..	..	..	..	..	..		..	..
04	..	..	..	..	..	..	..	..	..	..
05	..	..	13	16	17	..	..	..	..	..
06	..	..	9	12	14	16	..	..	..	..
07	..	..	..	8	11	14	15		..	..
08	..	..	..	9	·13	16	18	20	..	
09	..	..	..	8	12	16	18	18	..	
10	..	..	..	6	10	15	17	19	21	..
11	..	..	..	5	10	14	17	19	21	..
12	..	..	..	5	10	14	16	17	19	..
13	..	..	..	5	9	13	15	14	..	..
14	..	..	..	8	11	15	16	17	..	..
15	..	..	..	11	14	17	17	..	..	..
16	..	..	7	11	13	14	..	..	..	..
17	..	..	5	9	..	..	..	..	..	..
18	..	..	..	..	..	..	..	..	..	..
19	..	..	..	..	..	..	..	..	..	..
20	..	..	..	..	..	..	..	..	..	..
21	..	..	..	..	..	..	..	..	..	..
22	..	..	..	..	..	..	..	..	..	..
23	..	..	..	..	..	..	..	..	..	..

Table 3: This table shows the S/N value for a link whose parameters are indicated in the heading of this table.

ASAPS V2.23 Field Strength Prediction -------------------------

Circuit 1: P2 Distance: 622 KM Date: 15 Aug 1991
Tx: P2 41.45 19.95 Bearings: 277 92 T-index: 125
Rx: Roma 41.90 12.48 Path: Short Path TxAntenna: WH25
FSet: HF-ES2 Noise: -148 dBW/Hz RxAntenna: ISOTROPIC
Mode: 1E Power: 400W BandWidth: 3.0kHz
PROBABILITY

UT	2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0	11.0
00	..	..	..	..	..	..	..	..	..	..
01	..	..	..	..	..	..	..	..	..	..
02	..	..	..	..	..	..	..	..	..	..
03	..	..	..	..	..	..	..	..	..	..
04	..	..	..	..	..	..	..	..	..	..
05	..	..	99	99	50	..	..	..	..	..
06	..	..	99	99	99	99	..	..	..	..
07	..	..	..	99	99	99	84	..	..	..
08	..	..	..	99	99	99	99	50	..	..
09	..	..	..	99	99	99	99	87	..	..
10	..	..	..	99	99	99	99	99	14	..
11	..	..	..	99	99	99	99	99	19	..
12	..	..	..	99	99	99	99	99	14	..
13	..	..	..	99	99	99	99	81	..	..
14	..	..	..	99	99	99	99	35	..	..
15	..	..	..	99	99	99	76	..	..	..
16	..	..	99	99	99	79	..	..	..	..
17	..	..	99	99	..	..	..	..	..	..
18	..	..	..	..	..	..	..	..	..	..
19	..	..	..	..	..	..	..	..	..	..
20	..	..	..	..	..	..	..	..	..	..
21	..	..	..	..	..	..	..	..	..	..
22	..	..	..	..	..	..	..	..	..	..
23	..	..	..	..	..	..	..	..	..	..

Table 2: This table shows the values of Probability for the same link show in table 1.

ASAPS V2.23 Field Strength Prediction -------------------------

Circuit 1: P3 Distance: 623 KM Date: 15 Aug 1991
Tx: P3 37.90 17.60 Bearings: 317 133 T-index: 125
Rx: Roma 41.90 12.48 Path: Short Path TxAntenna: WH25
FSet: HF-ES2 Noise: -148 dBW/Hz RxAntenna: ISOTROPIC
Mode: 1E Power: 400W BandWidth: 3.0kHz
PROBABILITY

UT	2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0	11.0
00	..	..	..	..	..	..	..	..	..	..
01	..	..	..	..	..	..	..	..	..	..
02	..	..	..	..	..	..	..	..	..	..
03	..	..	..	..	..	..	..	..	..	..
04	..	..	..	..	..	..	..	..	..	..
05	..	..	99	99	17	..	..	..	..	..
06	..	..	99	99	99	86	..	..	..	..
07	..	..	..	99	99	99	84	..	..	..
08	..	..	..	99	99	99	99	50	..	..
09	..	..	..	99	99	99	99	87	..	..
10	..	..	..	99	99	99	99	99	19	..
11	..	..	..	99	99	99	99	99	19	..
12	..	..	..	99	99	99	99	99	19	..
13	..	..	..	99	99	99	99	87	..	..
14	..	..	..	99	99	99	99	50	..	..
15	..	..	..	99	99	99	84	..	..	..
16	..	..	99	99	99	86	..	..	..	..
17	..	..	99	99	..	..	..	..	..	..
18	..	..	..	..	..	..	..	..	..	..
19	..	..	..	..	..	..	..	..	..	..
20	..	..	..	..	..	..	..	..	..	..
21	..	..	..	..	..	..	..	..	..	..
22	..	..	..	..	..	..	..	..	..	..
23	..	..	..	..	..	..	..	..	..	..

Table 4: This table shows the values of Probability for the same link show in table 3.

Optimisation of the Radio Spectrum:
The Role of Computer Radio Prediction Packages

By

Adrian Graham BEng CEng MIEE

Director

ATDI Ltd.
15 Kingsland Court
Three Bridges Road
Crawley RH10 1HL
West Sussex
United Kingdom

It is recognised that a useful measurement factor in the assessment of spectrum management effectiveness is the number of users per square kilometre per MHz. Given that there is a requirement to maximise this factor, this paper examines computerised methods currently available to assist in the optimisation of spectrum usage. A methodology to extend this technology towards a more critically engineered solution is then examined.

Improved Spectral Efficiency - the requirement

The requirement for better methods of utilisation of the spectrum has been driven by the ever increasing demands of users in both the military and civil spheres. In military terms, the driving factors include telemetry for "smart" munitions and Unmanned Airborne Vehicles (UAVs), proliferation of sophisticated equipment such as battlefield surveillance radar's and, as a principal driver, the move towards "digitisation of the battlefield". All of this requires not only greater numbers of links, but also considerably increased bandwidth to cope with the flow of information. In the civil environment the proliferation of new service types and the increase in user uptake have both combined to reduce the available free spectrum. This is not a trend which is likely to reverse in the foreseeable future. As a result of the increased civil usage, in a time of relative peace, substantial pressure has been placed on military administrations to free up additional spectrum for civil use. All of this points to the necessity for efficient spectrum management, both in terms of long term allocation of spectrum and, particularly for the military users, a dynamic assignment process capable of maximising the use of allocated bands whilst minimising inter-system and inter-unit interference.

Traditional methods of frequency assignment, still widely used within NATO member nations, have tended to be based on paper rather than computers. When these systems of spectrum management were originally set up and developed, they were adequate to the task but they have inherent limitations that will prevent them from keeping up with the spectrum management requirements of the future. The principal limitations are that the assignments made will be sub-optimal because it is not possible to test all reasonable possibilities and also the fact that they are cumbersome and cannot react quickly enough to dynamic scenarios. In line with many other areas of human endeavour, the future is definitely computerised and, at least conceptually, integrated.

Computerisation of existing assignment methods offers tangible benefits such as increased accuracy and, because of the speed at which calculations can be performed, a "what-if" capability which allows far more possible solutions to be tested, most normally according to some established numerical algorithm, but the likely improvements due to this alone are limited because the key issues lie elsewhere. Of these, the main factor will be the flow of information about users, equipment and assignments to the spectrum managers and from the management process to the implementation process. The types of information required are shown in Figure 1.

For true flexibility and to meet the requirements of dynamic assignment at short notice, every spectrum manager must have accurate and *timely* information on current allocations and assignments in both the military and civil spheres, together with other environmental data, as in Figure 2, which expands upon the information flow requirements. The coloured arrows show the direction in which the information is sent. The dark green line denotes the military network being managed. The light green line denotes the information flow channels for the spectrum management process. This could be partially military and civil, dependent on the required connections and will certainly be spatially distributed, with some elements being located in headquarters or technical departments and others being held by other, non-military agencies. One point of note is that parts of the information flow actually travel through the network being managed. In the diagram, the assigned frequencies are disseminated via the network and, equally important, network information pertinent to the environmental circumstances as the users are passed back. In this way, it would be possible to determine actual conditions on the battlefield and to factor those into the assignment process. This is expanded upon later in the paper.

One of the most crucial areas of implementing a truly dynamic spectrum management system will be the prediction of radio performance, again in a timely manner. In the process outlined above, it is likely to be one of the most significant consumers of information as well as being a significant critical-path activity which will directly influence the timeliness of the overall system. The information overhead of such a system derives from the requirement to provide accurate information to the assignment algorithms in order to gain the most accurate solutions; any model is only as good as its inputs.

There are several main information requirement areas. One of these concerns the technical characteristics of all systems to be used in the operational environment, in the form of an equipment database. This must perforce include friendly, enemy and neutral emitters and receivers, as shown in Figure 3.

Unlike the non-communications electronic warfare situation, a fully populated database of communications assets does

Paper presented at the RTO IST Symposium on "Frequency Assignment, Sharing and Conservation in Systems (Aerospace)", held in Aalborg, Denmark, 5-7 October 1998, and published in RTO MP-13.

not yet exist; nor does the strategy and infrastructure necessary to undertake such a task. Given that the data capture required to develop such a database will be on a larger scale (in terms of the number of equipment types; after all, the complete EW database will form a *component* of an equipment/potential-interferer database), much thought and effort will have to be put into this activity. On the other hand, the fact that much of the relevant information is already available in the form of manufacturer's brochures will make data capture less hazardous than the activities that have had to be employed to populate the EW database.

The second information requirement, shown in Figure 4, is the provision of dynamic order-of-battle (ORBAT) information about friendly, enemy and neutral assets. This task is vital for communications and interference assessments, but it is a difficult one; not only does an accurate intelligence picture have to be incorporated, so does the civil communications infrastructure and, in many of the circumstances likely to be encountered by NATO, an understanding of the existence and operation of unregulated equipment used by others in the operational area, possibly including irregular forces or smugglers.
Again, the capture (by CIS update and direction finding (DF) of uncontrolled emissions), processing and timely dissemination of such data will assume non-trivial proportions and the methods employed to accomplish this must be carefully designed. The ORBAT database will not normally feature all of the data fields required for spectrum assignment or electronic warfare, given the specialist nature of these tasks, and therefore there must be a link-up between the two databases to allow the correct information to be used in the assignment process.
The third area is that of environmental data, as shown in Figure 5.
Atmospheric conditions in particular are continuously varying in terms of temperature and pressure, both at ground level and in the column of air above the ground. This is the transmission path for the radio signal, and because it changes continuously, so does the probability of establishing a reliable communications link between two points on the Earth. This is the main consideration in HF communications, but it is also important for VHF frequencies and above. A complicating factor is that the range at which a transmission will interfere with another user on the same channel will also vary in the same manner, which means that it is not possible to specify a generic maximum number of channels that will always be available within a given operational area; the number of channels that can be supported may increase or decrease with the environmental conditions. In the future, it may be desirable to plan communications intensive phases of operations to coincide with the best communications conditions, although clearly the enemy may have access to the same type of information and may therefore expect an attack at those times.

At present, most computer based radio prediction systems for VHF and above tend to base predictions on statistical meteorological information for those bands affected primarily by atmospheric conditions, rather than the prevalent conditions, because generally that information is not known either in advance or in a real-time sense. These methods do generally give workable solutions, but at the expense of requiring an additional safety margin to overcome the information shortfall. This does necessarily result in sub-optimal spectrum usage efficiency.

One possible way of improving the performance of these predictions is by dynamic free channel search, as used in some cellular systems. This scheme works well, but would be difficult to implement on a battlefield which has to provide spectrum for various communication systems as well as radar and other non-communication systems, each of which will have its own spectral characteristics and interference rejection capabilities. It also suffers in that it is not an actively planned system and therefore there is always the possibility that at some later stage there may be problems that will be difficult to resolve. To overcome this, it would be necessary to incorporate some planning capabilities into the management system. In order for this system to work effectively, there would have to be a way of presenting performance information about the existing network in order to project forward; communication links that were not established and environmental data would be two such measures.

Given the complex nature of the atmosphere, it is unlikely in the short to medium term that it will be possible to maintain a real-time, operational theatre-wide picture of small-scale variations such as rainshowers, fog or sandstorms, but by using meteorological data it should be possible to determine representative figures to be applied for all spectrum management for this region which will be accurate over the space of a few hours or days. If it becomes possible to incorporate this information directly into radio predictions, it will then be possible to reduce the amount of safety margin used to separate users. This release of safety margin can then inject more flexibility into the planning process and improve the efficiency of spectrum usage.

The only problem is that to accomplish this objective, it will be necessary to have a method of capturing environmental data on an ongoing basis, and for disseminating it to planners involved in the deployment.

Given all of the problems inherent with data capture and timely dissemination noted above, do computer predictions really have a large role to play in the optimisation of spectrum management? The answer is unquestionably yes, because of the benefits offered in terms of maximising the number of users per MHz per km^2.

To illustrate some of the benefits provided by computerised radio planning aids, it is possible to examine three individual aspects where improvements can be made. The first has a great use in dynamic allocation although is actually a non-dynamic variable; the terrain. Terrain shielding offers the ability to re-use frequencies for users separated by relief features, even if their spatial separation is small in terms of distance on the ground.

The classic example would be that of two groups of radio system users operating on the opposite sides of a large, steep hill, using V/UHF tactical radios. Although perhaps only separated by a few kilometres, the risk of interference even if they are re-using the same channels is very small. A traditionally based assignment process would not necessarily provide an accurate assessment of the degree of shielding and it would therefore be impossible to take advantage of this feature without running the risk of interference. If the deployments are static, an existing computer radio prediction system could be used to verify that the channels can be re-used, but if the users are moving there would be the risk of interference should they move into locations where the shielding reduces. A system that monitored the locations and intentions of the two groups and continually analysed potential interference would however be able to predict problems in advance, and take action to avoid them. This would allow the same channel to be re-used for as long as possible, and then a re-assignment to be made just before an interference problem occurs. For this scenario, atmospheric effects are negligible compared to the terrain effects, and so the primary dynamic feature is the movement of the groups. This could in theory be passed to the spectrum management system from a Command Information System (CIS) of the type currently coming into service in many NATO armed forces.

As an indication of the scale of benefits of this type of system could offer, it is possible to model the scenario on a computer prediction tool using randomly positioned deployments in very hilly terrain. This is the type of environment in which the greatest benefits are likely to be obtained, because of the terrain variability and hence a high degree of variability in the signal levels experienced near ground level. In simulations of this type, a small study indicated that compared to a traditional method (an exclusion zone 30 km around the base station; the radius was based on the field strength obtained after applying free space loss to the transmitted signal), the actual area where interference would be actually be a problem accounted for as little as 2 - 5 % of the excluded zone. The opportunity for channel re-use is therefore significant, but only if it can be managed appropriately. In less severe terrain environments, the level of gain was less but still significant, until the two methods converged on perfectly flat terrain. An example of a particular instance of this situation is shown in Figure 6. Further studies are required to determine how well this system could work in practice.
For airborne and rapidly moving objects it is more difficult to use terrain shielding, but this does not necessarily mean that computer-aided planning has nothing to offer. On the contrary, because of the dynamic nature of this scenario, computerised planning is the only realistic flexible real-time approach that is available. It is worth illustrating the possibilities of this approach by example.

Consider the scenario of an airborne element, perhaps carrying radio jamming equipment, moving through an operational area in which friendly forces are operating. Although the jamming equipment may be centred on a particular range of target frequencies used by the enemy, it is possible that out-of-band noise will affect a large number of channels over a wider range. If these channels are in use by ground forces, communications will be disrupted for the period of time during which the jammer is passing through. If the ground forces are to be able to communicate during this period, they must be given temporary assignments. This would only be possible with a dynamic frequency assignment system capable of actively modifying assigned channels whilst the users are in the field. A fully automatic system which is transparent to the users would most likely provide the best answer. This implies an intelligent network capable of automatically handling rapid changes without user intervention; we are a long way from this objective.

Again an essential part of this process would be the accurate simulation of the radio propagation throughout the battlefield at the time of the aircraft's transit. Ideally, this could be determined by some method of sensing the performance of network elements already operating within the area and passing the information back to an adaptive simulation process. Again, this requires a bandwidth and processing overhead on the communications links, but it is likely that the additional infrastructure that can be deployed will more than compensate for this.

The third area of improvement offered by computerised radio prediction methods is also concerned with the inclusion of more accurate and precise information into the spectrum management process, and the use of this information dynamically throughout the battlefield. This concerns the technical parameters of the transmitted signals and the way in which they are transmitted. At present the primary method of protecting against interference assuming that the location is fixed is to use directional antennas with a null, or close to it, in the direction of the interfering signal. In the future, more sophisticated active antennas may be able to generate more flexible responses to minimise interference by dynamically changing the antenna polar pattern in response to the requirement pertaining at the time. Digital systems using configurable modulation schemes could also be used to modify the characteristics of the transmitted signal, both to prevent unauthorised reception and, prospectively, to select the most robust scheme to overcome any interference problems detected. Although these are hardware solutions, they would provide the greatest benefit if combined with real-time dynamic management systems, which will use simulation to predict performance and to be the decision-making criteria.

The implementation of these techniques is dependent on introduction of new hardware and the subsequent generation of extensive technical characteristics (probably in the form of "look-up tables"). Again this task is likely to be non-trivial in nature, but the potential benefits are clear.

All of the methods outlined above offer potentially significant improvements in the areas of communications reliability, available link bandwidth, resilience to interference and, of course, maximising spectral efficiency. But increased reliance on technology always raises new risks as well, and this is particularly true in this case.

Whenever computer systems are used, there are always risks associated, including the effects of power disruption and the possibility of errors in the code, which in sophisticated systems may be far from obvious and have insidious effects. With the type of system outlined here, these risks are compounded both by the distributed nature of the system and the requirements for constant communications and dissemination of information throughout the network. This risk itself is again compounded by the possibility that the network being controlled by the management system is the only available transmission m edia.

Should the network fail for any reason, things will start to go awry very quickly, because the dynamic operational situation will continue to be dynamic, even if the network management scheme is non-operational. The first things to go wrong will be those that are changing the fastest; dynamic aircraft assignments and terrestrial allocations in response to aircraft movements. Next would be moving terrestrial units where terrain shielding has been used to inhibit interference and re-use channels. As these units move to locations where terrain shielding no longer works, the interference problems are compounded. And guess what? Because the only method of re-assigning channels is via the blocked channel, nothing can be done - unless of course a robust recovery mechanism has been designed in to the system.

The design of this recovery system is likely to be a difficult process, simply because the users in the field have come to depend on an optimal solution and by definition, any fallback option is almost by definition likely to be sub-optimal. The problems will stem from the fact that without sophisticated automated management processes, which require large amounts of information from users in the field (who are of course cut off by the system failure), there are simply too many users and too few resources available. A mechanism to build the whole process up again must be designed, and one which minimises the disruption during this phase, which is likely to take hours rather than seconds. A very complex system of user priority and traffic routing must be established; there are however precedents and the lessons learned from existing NATO command and control systems should prove invaluable; especially those hard-earned lessons gleaned during major exercise which many of us have experienced first hand.

From the outline analysis offered in this paper it can be seen that computer radio prediction packages are important to modern spectrum management, and their importance within the concept of automated network and spectrum management in the future will only continue to increase. The benefits offered will without question revolutionise the methods currently in use, but this improvement comes with associated risks which are real and which could potentially be devastating. Because of this that it will be necessary to carry out comprehensive studies to identify both the level of improvement the various possible implementations of such a system will offer, and the methods of implementation and management that will minimise the threat of system failure and its implications.

Whatever the future holds in terms of spectrum management and network management generally, it is clear that there is much work to be done, and for optimum efficiency, this work must be carried out in a co-ordinated across the NATO countries. The risks associated with modernisation are always significant, but the risks of failing to modernise are considerably worse.

Data Requirements for Dynamic Frequency Assignment

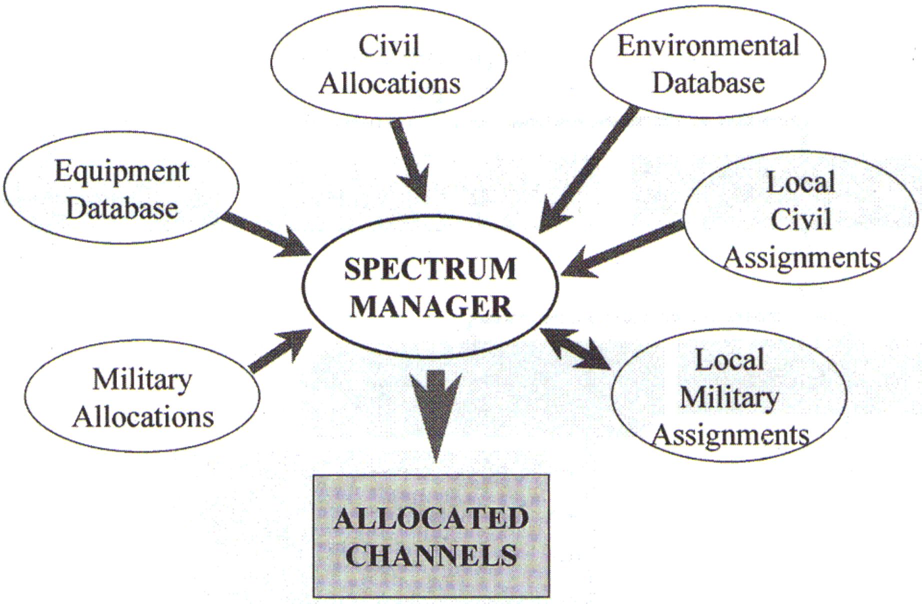


Figure 1: Data requirements & Flow

Figure 2: Information flow through a dynamic spectrum management system

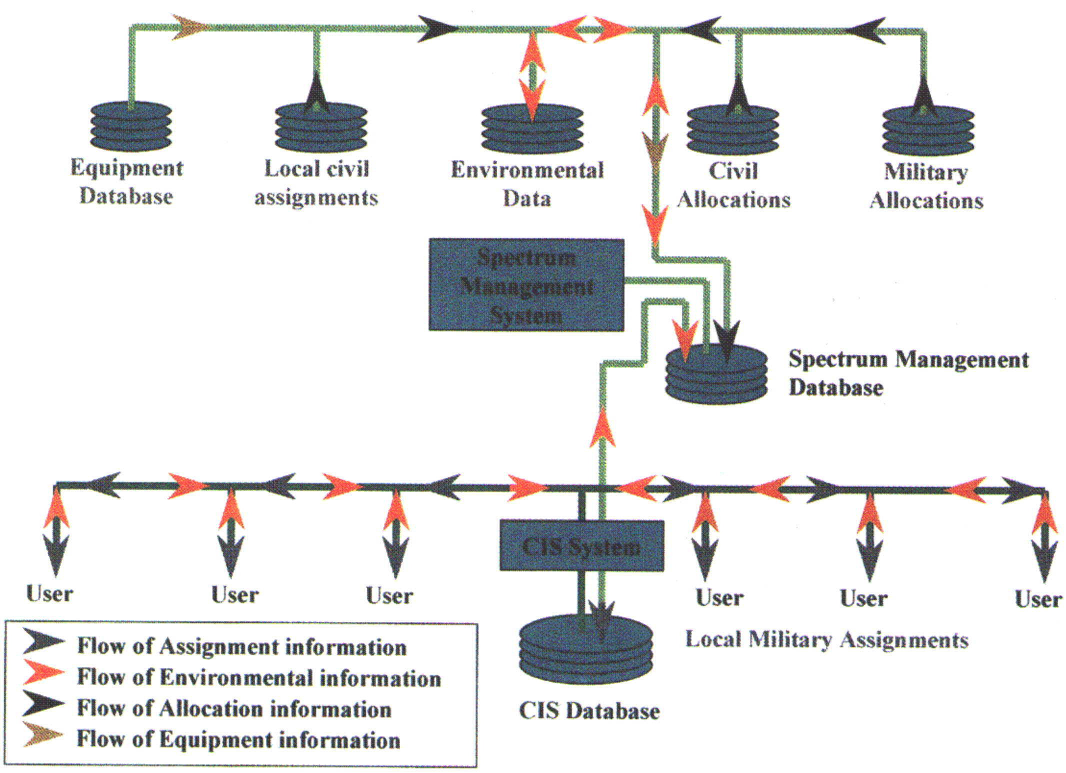


Figure 3: Equipment technical information requirements

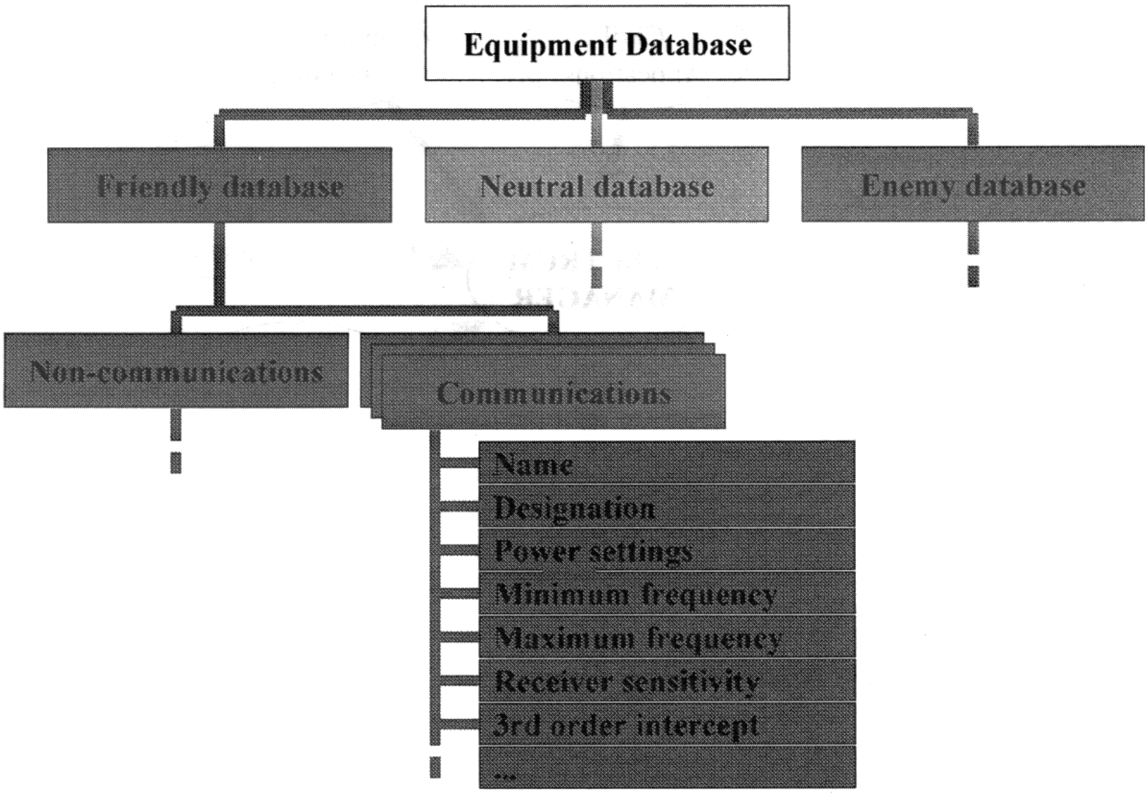


Figure 4: ORBAT information requirements

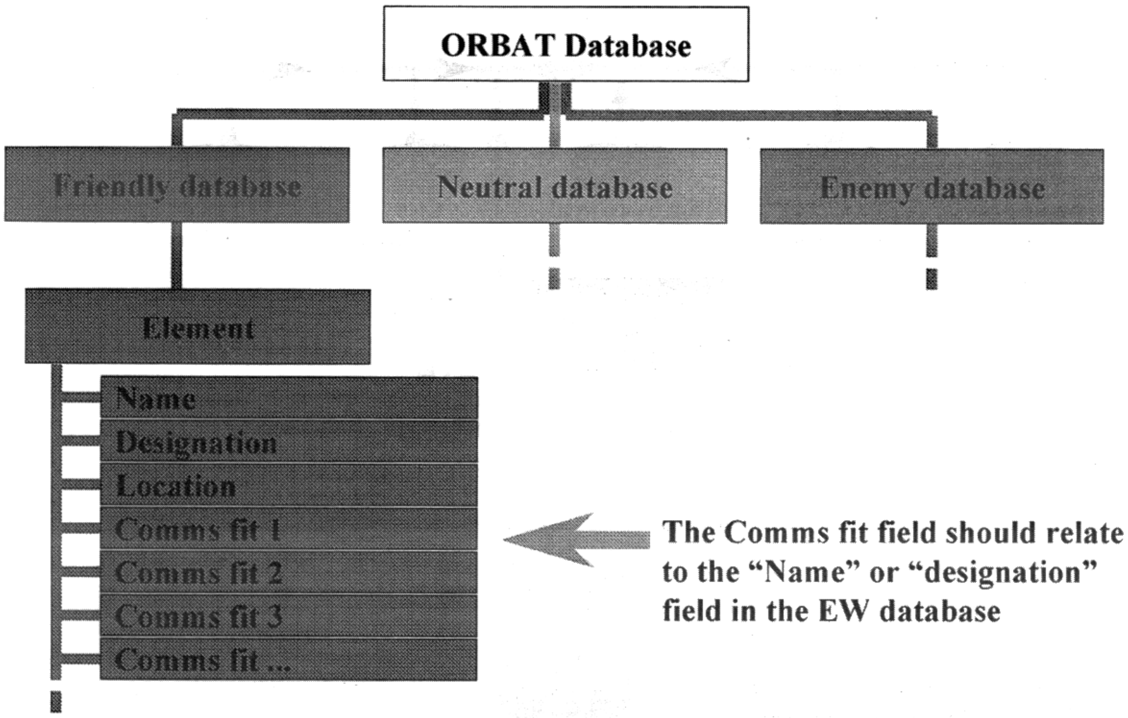


Figure 5: Environmental information requirements

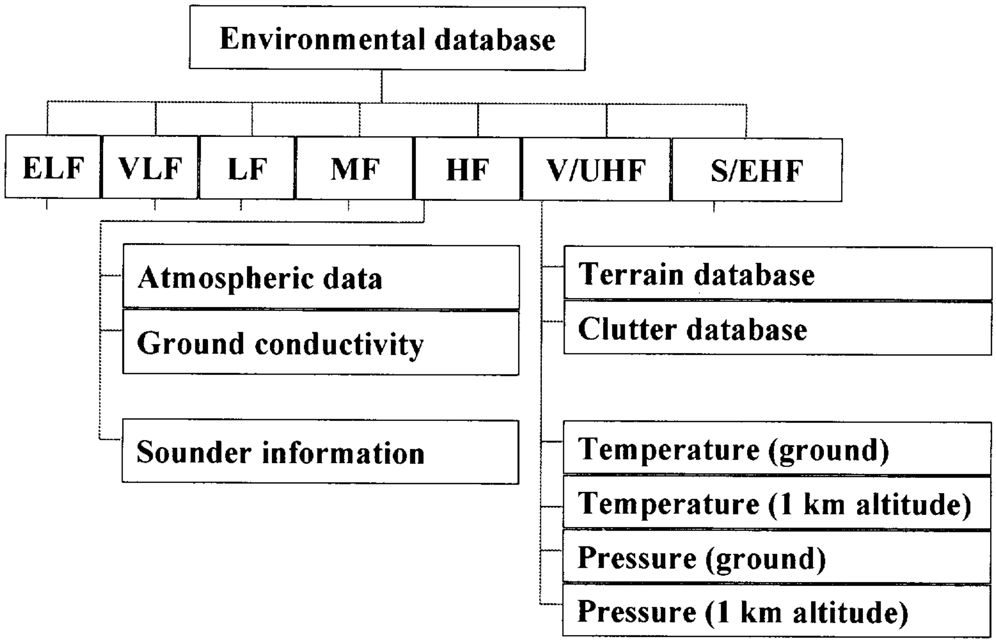
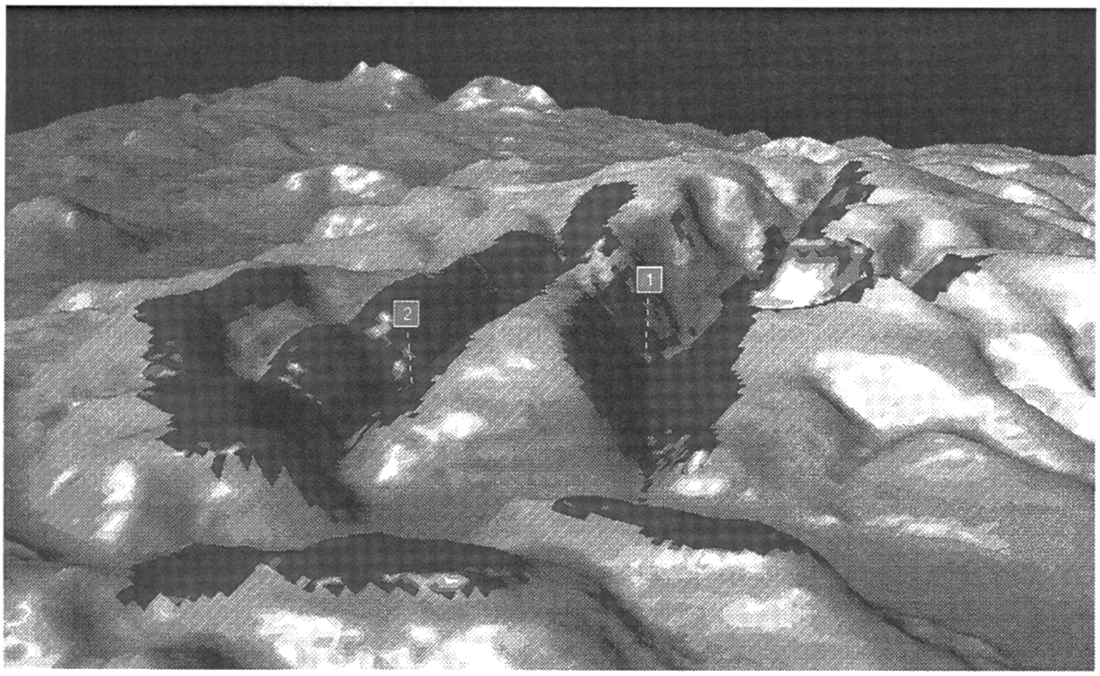


Figure 6: Terrain shielding



This image shows 2 groups of users sharing the same channel. The red overlay shows the limit of coverage to the minimum signal strength that will cause interference. Because of terrain shielding, no interference will occur.

PAPER TITLE : **Paper 5 - Optimisation of the Radio Spectrum: The Role of Computer based Radio Prediction Packages**

AUTHOR : A.W. Graham

NAME : R. Kapp-Rawnsley

QUESTION :

The techniques described will require high performance computer equipment which must be deployed into the field. Would you care to comment?

ANSWERS :

Our experience is that currently available PCs are suitable to the task. Because military planners are increasingly using sophisticated decision aids, such as Geographic Information Systems (GIS), they are also equipped with PCs that will be able to cope. Also, the experience they are gaining in the practical implications of using PCs in the field is highlighting potential problems.

NAME : G.R. Bradbeer

QUESTION :

As indicated, dynamic Spectrum management involves the increase of traffic over the networks supported for monitoring and control purposes. Can you comment on the extent to which the need for this traffic impacts on the benefits of dynamic spectrum management.

ANSWERS :

Clearly, this is a major issue for low bandwidth systems. It is crucial that any management system is appropriate and therefore, if the management process involves too high an overhead, it should not be used. This has to be assessed on a per-service basis.

PAPER TITLE : **Paper 5 - Optimisation of the Radio Spectrum: The Role of Computer based Radio Prediction Packages**

AUTHOR : A.W. Graham

NAME : K.S. Kho

QUESTION :

1. Which frequency bands are considered in your analysis?
2. Do you include spread spectrum in your analysis?
3. Changing wave form or frequency e.g. from fixed frequency to frequency hopping is not trivial. It is affecting a lot of users and also the frequencies should also be available/authorised. In a simulation model it may be possible but in practice we have very strict limitations in the usage of frequencies.

ANSWERS :

1. The analysis is intended to cover any type of radio service. The actual methods employed in a practical system will have to be appropriate to that service.

2. The management of spread spectrum systems is a difficult process, and it may be that this approach cannot be used, certainly for assessment of hopping systems. Practically, design considerations may offer the best management process.

3. Some management process of frequency hopping systems must be applied, although the exact method may be different from those applied to other services. In all cases of management, the method applied must be appropriate. As for the co-existence of fixed and frequency hopping systems, again there must be some process of management otherwise chaos results. The crucial issue is to identify a specific mechanism, and perhaps more work needs to be done.

Real-valued frequency assignment

D. G. Fon-Der-Flaass *
Centre for Discrete and Applied Mathematics
London School of Economics
Houghton St. London WC2A 2AE, U.K.

Abstract

We consider the binary constraints formulation of the frequency assignment problem in its most general form: for an arbitrary metric space, with frequencies taking arbitrary real values, and with possibly infinitely many constraints. We obtain some necessary and sufficient conditions for the problem to have a solution with a finite span. When the metric space is the set of integers, we give an exact criterion.

Also we demonstrate a connection of this problem in one-dimensional case with one combinatorial question about finite permutations; and pose some unsolved problems.

The practical problem of assigning frequencies to transmitters in a cellular network has been approximated by various combinatorial and/or geometrical formulations. Most often it is treated as a problem of finding a colouring of a metric space (e.g. a graph). The required colouring should satisfy a certain set of restrictions; usually each restriction involves two vertices, and the restrictions therefore are called *binary constraints.*

In this note we consider a variant of the standard binary constraints problem in which frequencies (or *colours*) can take on arbitrary real values. Apart from its theoretical significance, this problem can also have some practical importance; especially for the question to what extent one can improve the performance by using more channels within the same frequency interval.

Let (M, d) be an arbitrary metric space. By a (real-valued) *colouring* of M we mean any function $c : M \to \mathbf{R}$; the *span* of the colouring being

$$sp(c) = \sup_{m \in M} c(m) - \inf_{m \in M} c(m).$$

By a *restriction function* we mean any non-increasing function $f : \mathbf{R}^+ \to \mathbf{R}^{\geq 0}$. A colouring c of M *satisfies* the restriction f if for every two distinct elements x, y of M we have

$$|c(x) - c(y)| \geq f(d(x, y)).$$

The real-valued frequency assignment problem in its most general form can now be stated as follows:

given a metric space M and a restriction function f, find the value of inf $sp(c)$ for all colourings c of M satisfying the restriction f.

We shall denote this number by $sp(M, f)$.

First we shall study this problem for $M = \mathbf{Z}$, the set of integers, with the natural metric $d(x, y) = |x - y|$. The restriction function in this case is determined by a non-increasing sequence $f = (f_n)_{n>0}$ of non-negative real numbers. The first question which arises in this situation is whether $sp(\mathbf{Z}, f)$ is finite or infinite. We shall give the complete answer to this question, and also we'll find upper bounds on the span when it is finite.

THEOREM 1 *For the restrictions sequence $f_n = 1/n$, the span $sp(\mathbf{Z}, f)$ is finite and does not exceed*

$$1 + \varphi = 2.618\ldots,$$

where $\varphi = (1 + \sqrt{5})/2$ is the golden ratio. There exists a colouring c of span $1 + \varphi$ satisfying f.

THEOREM 2 *For an arbitrary restrictions sequence $f = (f_n)$, the span $sp(\mathbf{Z}, f)$ is finite*

*Supported by the U.K. Radiocommunications Agency

if and only if $C = \sup\{nf_n \mid n > 0\}$ *is finite; and the span does not exceed* $C(1+\varphi)$.

THEOREM 3 *If the sum* $s = \sum_{n>0} f_n$ *is finite then the span* $sp(\mathbf{Z}, f)$ *is finite, and the span does not exceed* $2s$.

First we shall derive Theorem 2 from Theorem 1, and prove Theorem 3; both proofs are fairly easy. Then we'll give a proof of Theorem 1.

PROOF OF THEOREM 2. The "if" part immediately follows from Theorem 1. Indeed, if the number C is finite then for all n, $f_n \leq C/n$; and we can take the colouring c from Theorem 1 multiplied by C.

The "only if" part is just as easy. Given $L > 0$, choose n such that $nf_n > L$. As $f_i \geq f_n$ for $i < n$, all differences between colours of elements $0, 1, \ldots, n$ in a colouring satisfying f are at least f_n; and the span of these $n+1$ colours is at least $nf_n > L$. □

PROOF OF THEOREM 3. We shall assign colours to elements of $\mathbf{Z}$ one by one, in the following order: $0, 1, -1, 2, -2, \ldots$, choosing on each step a value $c(k) \in [0, 2s]$ which satisfies the colouring conditions. This is always possible, since the measure of the set of values for $c(k)$ which are forbidden by any of the previously chosen values of $c(i)$, $i \leq |k|$, is at most $2\sum_{i=1}^{2|k|} f_i$ and is less than $2s$. □

PROOF OF THEOREM 1.

Denote by $a \bmod b$ with real a, b the number q such that $0 \leq q < b$, and $(a-q)/b \in \mathbf{Z}$. Let also $\mathbf{R}_r$ mean the factorgroup $\mathbf{R}/r\mathbf{Z}$ of the additive group of reals. The distance between elements of the group is naturally defined as $d_r(a, b) = \min((a-b) \bmod r, (b-a) \bmod r)$.

Let X be the half-open interval $[-(\varphi+1)/2, (\varphi+1)/2)$. Define a colouring $c : \mathbf{Z} \to X$ by the formula: $c(n) = ((\varphi+1)/2) + n\varphi) \bmod(\varphi+1) - (\varphi+1)/2$. If we consider elements of X as representatives of the cosets $\mathbf{R}/(\varphi+1)\mathbf{Z}$ then the colouring c becomes a group homomorphism $c : \mathbf{Z} \to \mathbf{R}_{\varphi+1}$.

Let us prove that this colouring satisfies the restriction f. Since c is a homomorphism, it is enough to show that for every $n > 0$ the distance $d_{\varphi+1}(c(0), c(n)) \geq 1/n$.

Our colouring has the following easy properties:

(*i*) $c(0) = 0$;

(*ii*) $c(n+1) - c(n)$ is equal to either φ or -1;

(*iii*) $d_{\varphi+1}(0, c(n)) = |c(n)|$.

The property (*ii*) enables us to give an alternative description of the colouring by an easy inductive process: $c(0) = 0$, and $c(n+1)$ is that one of the two numbers $c(n)+\varphi$, $c(n)-1$ which belongs to X.

Properties (*i*) and (*ii*) imply that the number $c(n)$ is of the form $-a+b\varphi$ for some natural numbers a, b with $a+b = n$. Let $y = b + a\varphi$; we have $y \leq n\varphi$. Also,

$$c(n)y = ab(\varphi^2-1)+(b^2-a^2)\varphi = (ab+b^2-a^2)\varphi.$$

The number $ab+b^2-a^2$ is a non-zero integer; therefore $|c(n)| \geq \varphi/y \geq \varphi/(n\varphi) = 1/n$, and by (*iii*) we are done. □

Note that the colouring described above is a cyclic channel colouring (in the sense of [2]) with the same constraints function; which is a stronger property.

Our next theorem gives a sufficient and a necessary condition for finiteness of the span for general discrete metric spaces. Though these conditions don't give a complete answer to the problem, they are rather close.

Let M be a metric space, and sequences (l_n), (u_n) be such that for every $n > 0$, and for every $x \in M$, the following inequalities hold:

$$l_n \leq |\{y \in M \mid n-1 < d(y, x) \leq n\}| \leq u_n.$$

Let f be a restriction function; we extend its domain by setting $f(0) = \sup_{x>0} f(x)$, and assume that $f(0)$ is finite. This assumption is very natural. Indeed, if there is a positive lower bound ε on the distance between distinct elements of M then we can assume that $f(x) = f(\varepsilon)$ for all $0 < x < \varepsilon$; on the other hand, if one can have arbitrarily small distances then this assumption is necessary for the span to be finite.

THEOREM 4 *In the above notation, if* $\sup f(n) \sum_{i=1}^{n} l_i = \infty$ *then* $sp(M, f)$ *is infinite. If* $\sum_{i=1}^{\infty} f(i-1)u_i < \infty$ *then* $sp(M, f)$ *is finite.*

PROOF. The first assertion is proved exactly as the "only if" part of Theorem 2; the second assertion — as Theorem 3. □

For the k-dimensional rectangular grid $\mathbf{Z}^k$ we can take $l_n = c_1 n^{k-1}$ and $u_n = c_2 n^{k-1}$ for suitable constants $0 < c_1 < c_2$. As an immediate corollary, we get that $sp(\mathbf{Z}^k, (n^{-\alpha}))$ is finite when $\alpha > k$, and infinite when $0 < \alpha < k$; and this holds for every metrics on $\mathbf{Z}^k$ equivalent to the Euclidean one — including the graph distance metrics.

The question whether $sp(\mathbf{Z}^k, (n^{-k}))$ is finite or infinite remains open for $k \geq 2$.

The result of Theorem 1 has an interesting application to monotone subsequences in permutations.

The classical theorem of Erdős and Szekeres [1] states that every permutation of more than n^2 numbers has a monotonic (increasing or decreasing) subsequence of length more than n.

Let $s \in S_n$ be a permutation of the set $\Omega = \{1, \ldots, n\}$, and $w = (i_0, \ldots, i_k)$ be an increasing sequence of elements of Ω such that the sequence $(s(i_0), \ldots, s(i_k))$ is monotonic (to avoid trivialities, we assume that $k \geq 2$). By the *density* of w we mean the number $d(w) = k^2/(i_k - i_0)$.

The above-mentioned theorem implies that for $n > 4$ there exists a lower bound $b(n) > 0$ such that every permutation in S_n has a monotonic subsequence of density at least $b(n)$; the bound $b(n)$ so obtained tends to 1 as n tends to infinity. A question arises: can one find such a lower bound which would tend to infinity with n? Theorem 1 gives a negative answer to this question.

Take the colouring c as in this theorem. For every n we define a permutation $f_n \in S_n$ as follows. The numbers $c(1), \ldots, c(n)$ are all distinct and linearly ordered; let $f_n(k)$ be the position of $c(k)$ in this ordering. We claim that every monotonic subsequence of any of the permutations f_n has density less than $1 + \varphi$. Indeed, let $i_0 < i_1 < \ldots < i_k$ be such that the sequence $(c(i_0), \ldots, c(i_k))$ is monotonic. We have:

$$\begin{aligned} i_k - i_0 &= (i_k - i_{k-1}) + \ldots + (i_1 - i_0) \geq \\ &\geq |c(i_k) - c(i_{k-1})|^{-1} + \ldots + \\ &\quad + |c(i_1) - c(i_0)|^{-1} \geq \\ &\geq k \cdot (|c(i_k) - c(i_0)|/k)^{-1} > \\ &> k^2/(1+\varphi); \end{aligned}$$

and the claim follows.

This leaves two related open questions: find the best possible value b_1 of the bound b from the above problem; and find the exact value of $b_0 = sp(\mathbf{Z}, (1/n))$. The above considerations show that

$$1 \leq b_1 \leq b_0 \leq 1 + \varphi = 2.618\ldots.$$

It is not difficult to show that $b_0 \geq 2$ by considering all possible orderings of the five values $c(0), \ldots, c(4)$.

References

[1] P. Erdős, G. Szekeres. A combinatorial problem in geometry. *Compositio Math.* 2(1935), 463–470.

[2] J. van den Heuvel, R.A. Leese, M.A. Shepherd. Graph Labelling and Radio Channel Assignment. *CDAM Research Report Series*, CDAM-96-23, 1996.

PAPER TITLE : Paper 6 - Real-Valued Frequency Assignment

AUTHOR : D.G. Fon-Der-Flaass

NAME : E. Tsang

QUESTION :

Are you suggesting that all binary graph colouring problems are easy?

Comment: Research from the constraints programming community finds that the majority of tandem graph colouring problems are easy to solve in practice, (although the problem is NP-complete in nature), with some exceptions.

ANSWER :

The graph colouring problems arising in frequency assignment are "easy" only in the sense that the span is within a constant factor of lower bounds obtained from cliques, as I have demonstrated. For arbitrary random graphs, which do not arise from a placement of transmitters on the plane, this is not so. This partly explains why meta-heuristic algorithms work so well for typical frequency assignment problems, but fail for random graphs.

Paper 7: HF Frequency Management: Prediction and Assignment Tools for Large Networks.

By
N.S.Wheadon
GEC-Marconi Research Centre, West Hanningfield Road, Great Baddow, Chelmsford, Essex, CM2 8HN, UK.

nigel.wheadon@gecm.com

Introduction.

With the advent of automatic and adaptive radio communications systems there has been a resurgence of activity in propagation at HF.

The development of Automatic Link Establishment (ALE) and Automatic Link Maintenance (ALM) systems suggested the demise for HF prediction programs. These new systems would not need experienced operators and the systems would perform all the necessary frequency management tasks themselves. ALE and ALM have gone some way in optimising the operation of systems on air; however in order to optimise the use of the HF spectrum, a planning exercise is still required which will provide a reasonable first selection of frequencies. For example, allocating the high end of the band for short range near-vertical incidence sky-wave operations in winter at midnight would not necessarily be a sensible choice. This "first cut" planning exercise could be performed using radios which contain simple software planning tools but co-ordination between separated users wishing to communicate becomes an issue. Alternatively the planning exercise can be performed using more sophisticated assignment tools located at a strategic centre or at a tactical command post.

This paper describes two tools which can be used to perform tasks from the simple estimation of usable frequencies to more complex frequency allocation for large networks.

Frequency Planning Tools

Two software planning aids will be described. The first tool, WinHF, is a combined ground-wave and sky-wave propagation prediction program used by the UK MoD across all three services. The second, FAA, is a sophisticated and complex tool for generating and evaluating frequency assignments for very large systems (figure 1) comprising more than 1000 individual networks each consisting of many radios with the primary aim to minimise interference. After a description of the tools, a short demonstration will be performed highlighting the main features of each program; in the time available each demonstration will only be able to show a small subset of the features implemented.

WinHF

For several decades the Marconi Electronic Systems company has been involved with the development of HF propagation prediction models and has provided them to Government and Civil organisations. The predictions have evolved from laborious paper exercises, through the use of slide rule methods and books of graphical predictions to computer algorithms. There are a plethora of algorithms to choose from and until recently radio

Paper presented at the RTO IST Symposium on "Frequency Assignment, Sharing and Conservation in Systems (Aerospace)", held in Aalborg, Denmark, 5-7 October 1998, and published in RTO MP-13.

operators within the UK MoD would typically use several different programs in their HF communications planning phases and take a "majority" vote as to which frequencies they would use. Other radio operators would tend to stick with the frequencies which had proven to work in their experience. These various methods have been superseded by the Marconi WinHF program which brings together the users' operation requirement into one application. The new software package makes use of validated HF prediction methods.

WinHF provides the capability to make predictions for both sky-wave and ground-wave point-to-point circuits. It provides the user with a recommended window of frequencies using a graphical user interface (GUI) from which frequencies may be selected. The GUI presentation for the sky-wave and ground-wave parts of the program are different reflecting the different user requirements. For the sky-wave part, a diurnal frequency plot is provided which displays the envelop of usable frequencies together with a recommended frequency of operation defined as the frequency which gives the highest signal-to-noise ratio (figure 2). For the ground-wave prediction two graphical displays are provided (figures 3 & 4) enabling area coverage estimations and operational frequency trade-offs to be made.

WinHF Prediction Methods

The algorithm used for the ground-wave predictions is known as the Bremmer method which provides equivalent results to the ITU-R GRWAVE programme (also developed at GEC-Marconi in the 1980's) and is much faster in execution.

The sky-wave element is provided by the ITU-R Report 894 method. The ITU-R method has been validated by comparison of the field strength predictions against the ITU-R database of radio circuit measurements. The sky-wave program uses the ITU-R Report 340 ionospheric maps to perform the propagation calculations.

The noise field strength calculation uses the methods of ITU-R Report 322 for atmospheric noise determination and ITU-R Report 258 for the man-made and galactic noise characteristics.

Customisation

WinHF has been developed so that it is extremely flexible and easy to use and also for this version. it has been 'customised' with UK military radio and antenna equipment. Where possible the entry of numerical values from the keyboard has been removed thereby ensuring less 'finger trouble' problems. Even the sunspot number entry has been removed and entry is based on the selection of month and year - the sunspot number data being related to these two quantities by the program. To support these changes there is a database of parameter values which are under the configuration and change control and up dates to this database occur on a regular basis. Location co-ordinates can either be entered from the keyboard directly in a variety of formats or via the use of the locator map.

Overview of the Frequency Assignment Process

The majority of the work undertaken at the Research Centre is applied research. The development of an HF frequency assignment algorithm (FAA) is one such project. In fact the FAA is just one part, but a very important part, of a project which culminated in the development of a general HF Frequency Management Tool (FMT) with our sister company Marconi Communications. The FMT is used to generate complete frequency plans for large (and small) networks and it provides assignments for strategic communications planning (COMPLAN). Several different COMPLANs can be generated for different operational scenarios and the COMPLAN can be modified easily on a day-to-day basis to take account of changes within the network configuration. The FMT can also be used in the civil environment to manage the spectrum, for keeping records of licensed frequencies and for checking new assignments for interference effects with existing systems.

Components

The FMT comprises the FAA module which will be described in more detail below but essentially computes the frequencies which are to be assigned to each network and checks that the assignment is valid. The other key modules that the Research Centre provided for the FMT were

i) the sky-wave prediction algorithm
ii) the ground-wave prediction module
iii) the noise field strength prediction module

These three modules use the same methods as the WinHF software as described earlier in this paper.

The assignment methods make provision for the different types of networks e.g. the FMT has to assign frequencies for networks ranging from a few kilometres in extent to large networks (a few thousand kilometres and even world-wide), land, maritime and aero-mobile network configurations as well as point-to-point links. The networks may operate in either sky-wave or ground-wave modes. The FMT also allows for inter-working with broadcast stations. The FMT also allows for the inclusion and evaluation of external interferers and jammers within the frequency plans, so that these frequencies may be avoided.

The code is portable and runs under Unix or Windows NT platforms, and all the radio equipment parameters (i.e. co-channel bandwidth, adjacent channel bandwidth, interference protection ratios) are fully adjustable.

Databases

The frequency management process initially requires the user to generate a complete network structure. This activity comprises defining radio sites and mobile network locations and sizes, defining the required connectivity, the maintenance of an equipment inventory (e.g. radio types including serial numbers, antenna types and characteristics) and the allocation of those items to locations and networks, defining allocated frequency bands, unusable frequency bands, and so on. Figure 5 shows a simplified overview of the relationships between the networks, stations and equipment. The radio equipment is defined with the following parameters: Power output, harmonic radiation, transmit spurious radiation, AMU losses, band-pass filter rejection, transmit broadband noise, receiver sensitivity, co-site filter rejection and the list of associated antennas which can be used with each radio item. The antenna items have parameters including gain, feeder loss, polarisation etc.

Each network also requires parameters to describe its channel type e.g. single or multiple frequencies, simplex, duplex, etc. The emission classes are user defined but typically the list may include SSB, FSK, DSB, AM, etc. The radio equipment parameters such as co-channel bandwidth, adjacent channel bandwidth, interference protection ratios, etc., are fully adjustable for each network.

Assignment Updates

As the networks, and their configurations change, then it is a relatively simple task to modify the entries that have changed and re-run the FMT for re-generation of assignments. Various levels of re-run can be specified depending on whether a new frequency assignment is required, or to provide an assignment for a single new network. Within the networks specifications dialogue, the user can specify: mandatory assignments; preferred assignments which will be used if possible or with varying degrees of conflict with the existing networks as specified, or completely new assignments can be generated.

Assignment Phase

The assignment method comprises several distinct phases. The overall method can be described as a sequential algorithm, in terms of finding a good assignment to begin with, followed by a meta-heuristic technique which then tries to minimise the cost, or penalty due to interference with other networks, even further.

i) The generation of a list of usable parts of the spectrum. The final list of frequencies will typically exclude distress frequencies, time signals and frequencies which may be barred from use at a particular station e.g. a station may be located in a different ITU sector.
ii) The generation of an 'active frequency range' for each network i.e. identifying those parts of the spectrum were there is an adequate signal-to-noise ratio for information (data or voice) transfer (figure 6).
iii) The generation of an 'active frequency pool' for each network. The pool comprises a list of individual frequencies or multiple sets of frequencies for frequency hopping or ALE networks. The FAA then assesses the interference ('cost') between networks using frequencies from the pool.
iv) Networks are sequenced to facilitate an optimum order in which the networks are assigned. The reason for ordering the networks is to ensure that those networks that have specific requests for frequencies or have 'high' importance e.g. strategic communications links, are serviced first and their frequencies can then be protected during subsequent assignments.
v) Several frequencies, or frequency sets, are selected from the active frequency pool and then they are analysed to determine the 'best' assignment; the 'best' being the one which: exceeds the various signal-to-noise criteria, causes least interference and is also least affected by other networks (figure 7). The acceptance of a frequency, as well as a complete assignment for a network, comprises several assessment processes. The three processes are:
    a) The choice of frequency must be such that it causes minimal interference with the existing networks, unless otherwise specified. Both

co-channel and adjacent channel interference can be analysed.

b) There are several thresholds of acceptability for both co- and adjacent channel interference and these can be modified by the user for each network separately.

c) The computation of co-site effects can be enabled to ensure that the assigned transmit frequencies on a transceiver site do not give rise to intermodulation products on frequencies being used for reception purposes in the vicinity.

v) Once assignments have been made to all networks it is possible to force the assignment process to search or 'hunt' for a better assignment for both individual and multiple networks if required.

vi) Networks can also be assigned sets of back-up or 'reserve' frequencies if required.

Output

The output from the FAA is in the form of lists (electronic or paper copies) of frequencies for each of the networks together with factors which describe the quality of the assignment. Within the FMT then there are various GUI displays to enhance the output for visualisation of the assignments. For the radio end-user a down-loadable electronic list of frequencies is available for porting to an intelligent radio via a fill-gun.

FAA Demonstration

For the purposes of the demonstration that follows, a demonstration version of the FAA (rather than the complete FMT) will be shown. The FAA functionality described above is included in the version which will be shown. The network signal and interference calculations have been determined previously to save demonstration time. Only a small sub-set of the PC GUI is provided together with a simple output GUI, this is because the focus of the work at the Marconi Research Centre is on the algorithms rather than the complete man-machine interface and it is the former which we to emphasis.

Summary

Marconi Electronics Systems has over several decades developed HF frequency prediction tools and assignment algorithms. The WinHF program is in-use with the MoD providing them with an HF ground- and sky-wave communication prediction capability. The assignment algorithms and methods have matured and are now incorporated into various Marconi frequency assignment products

This paper and associated presentation has emphasised the radio-wave propagation and frequency assignment aspects of the new tools.

Acknowledgements

The development of the FAA was supported by GMCL, New Street, Chelmsford. The WinHF program development has been performed in-house and is now in use with the UK Armed Forces supported by the Defence Communications Procedures Branch, MoD, UK.

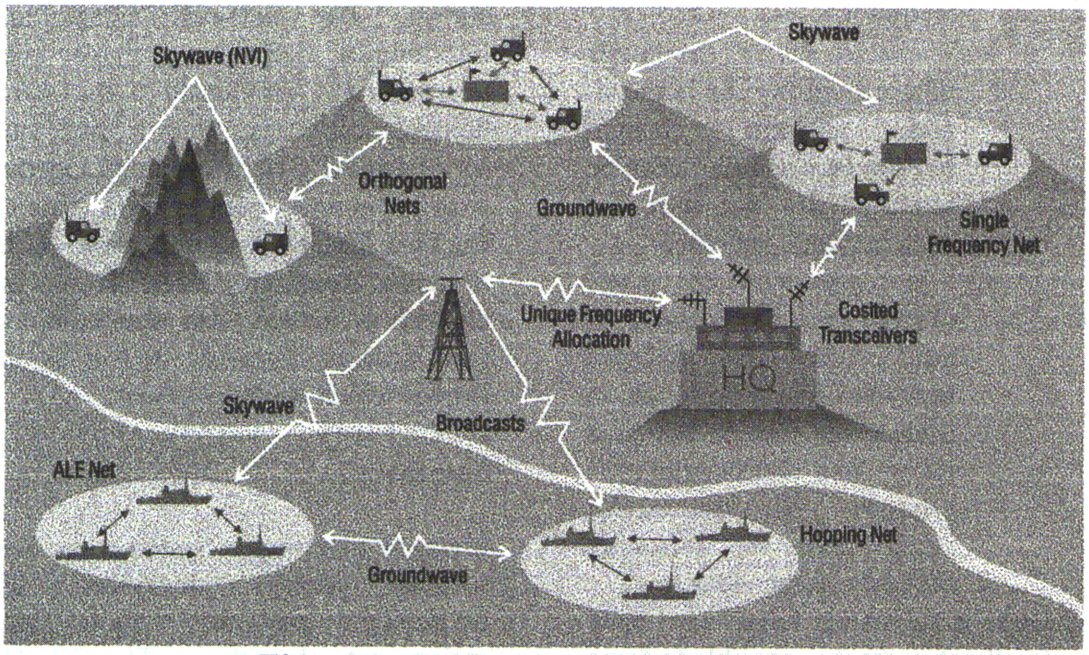
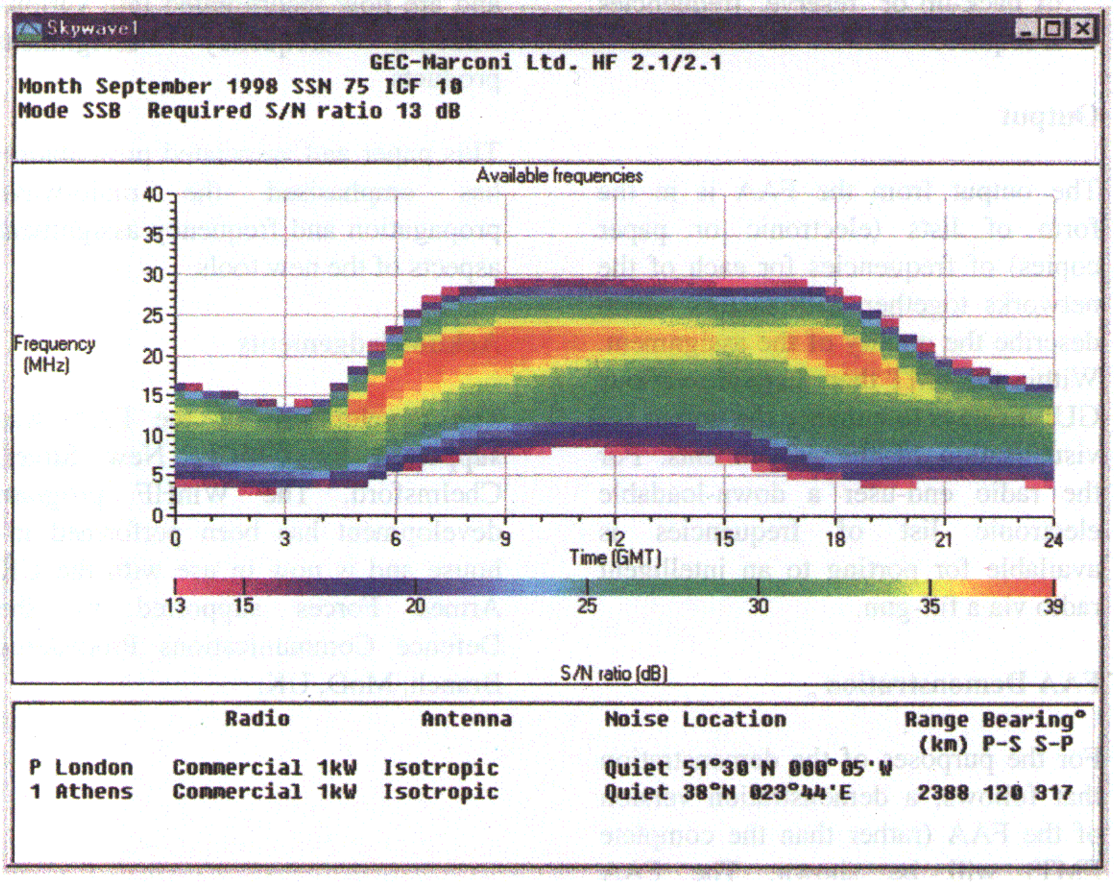


Figure 1. Typical large HF network.



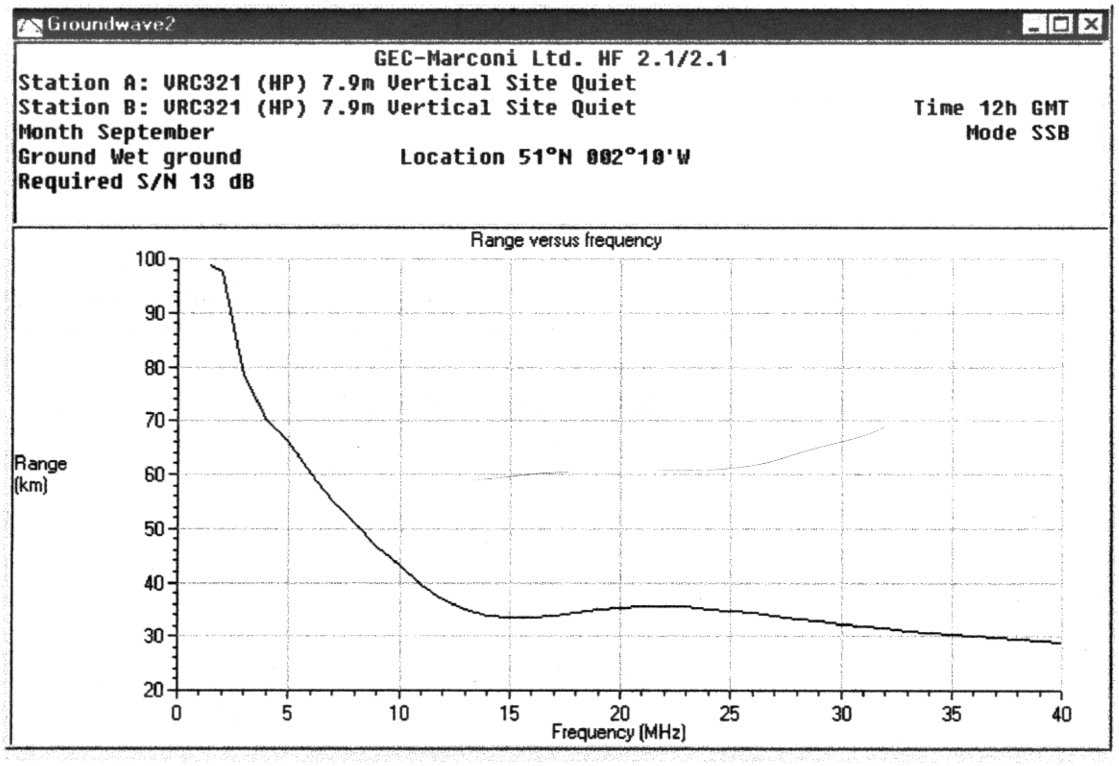


Figure 4. Ground-wave HF range v frequency plot.

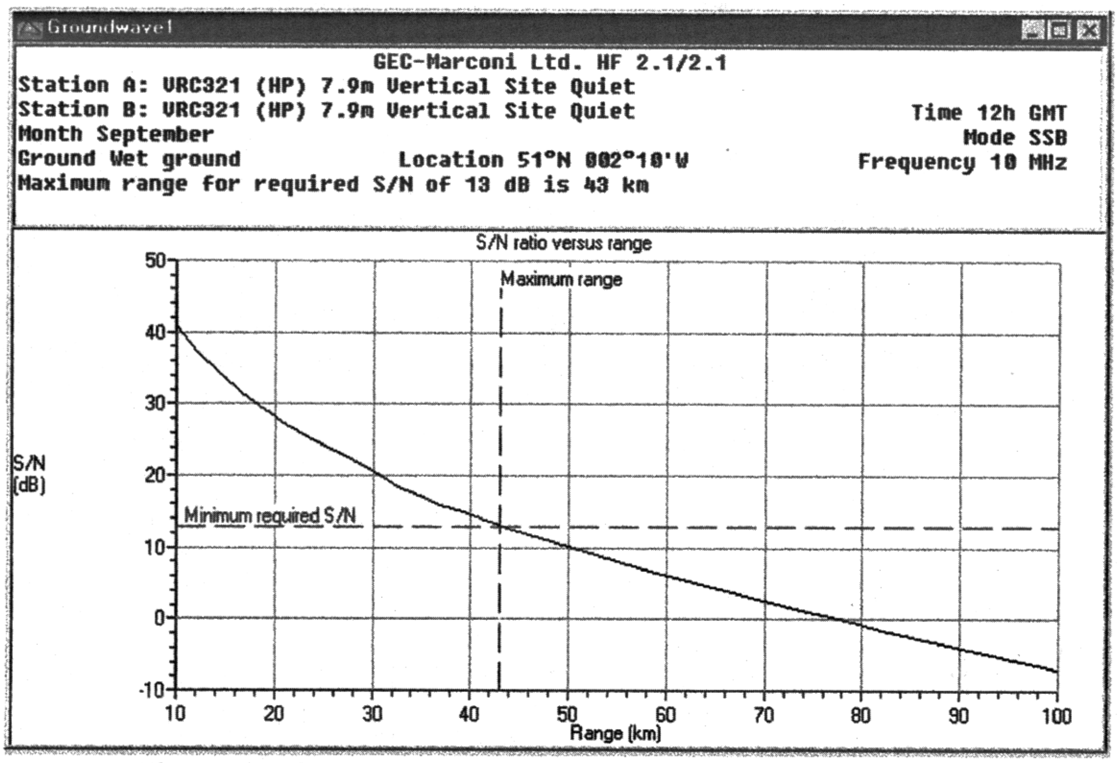


Figure 3. Ground-wave HF signal-to-noise vs range plot.

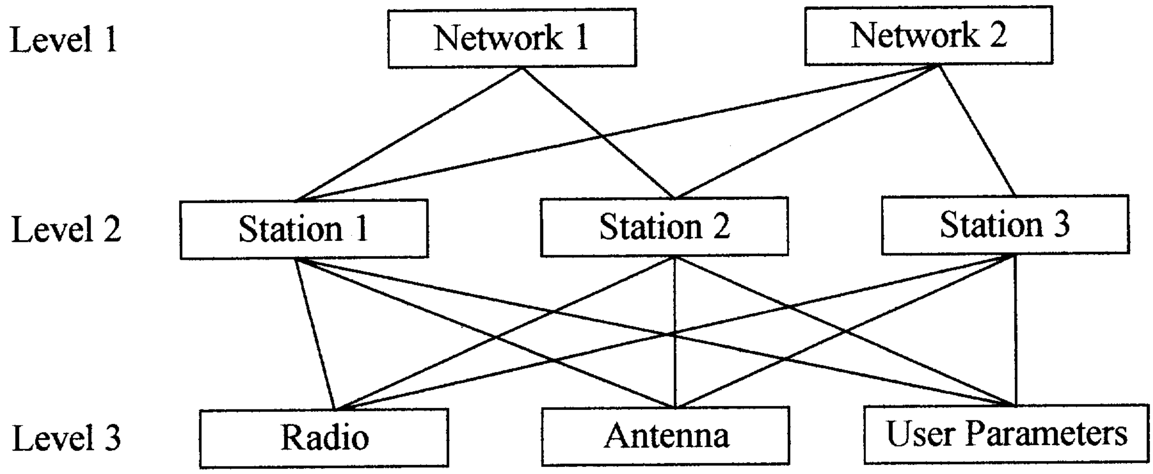


Figure 5. Network configuration and association.

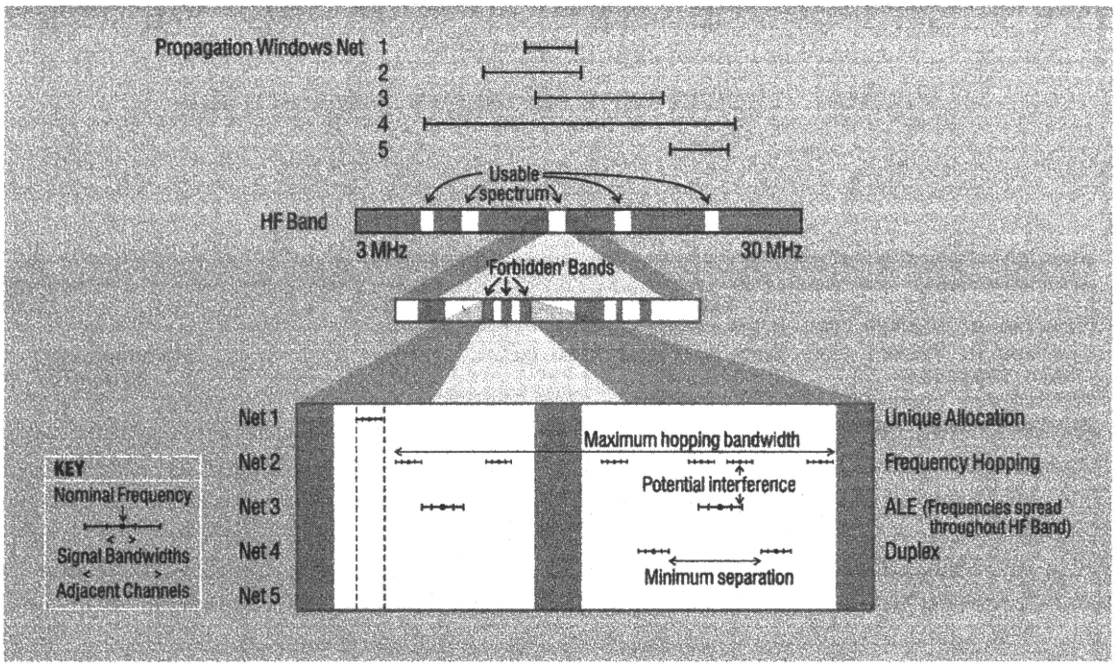


Figure 6. Active frequency pool and initial assignment process.

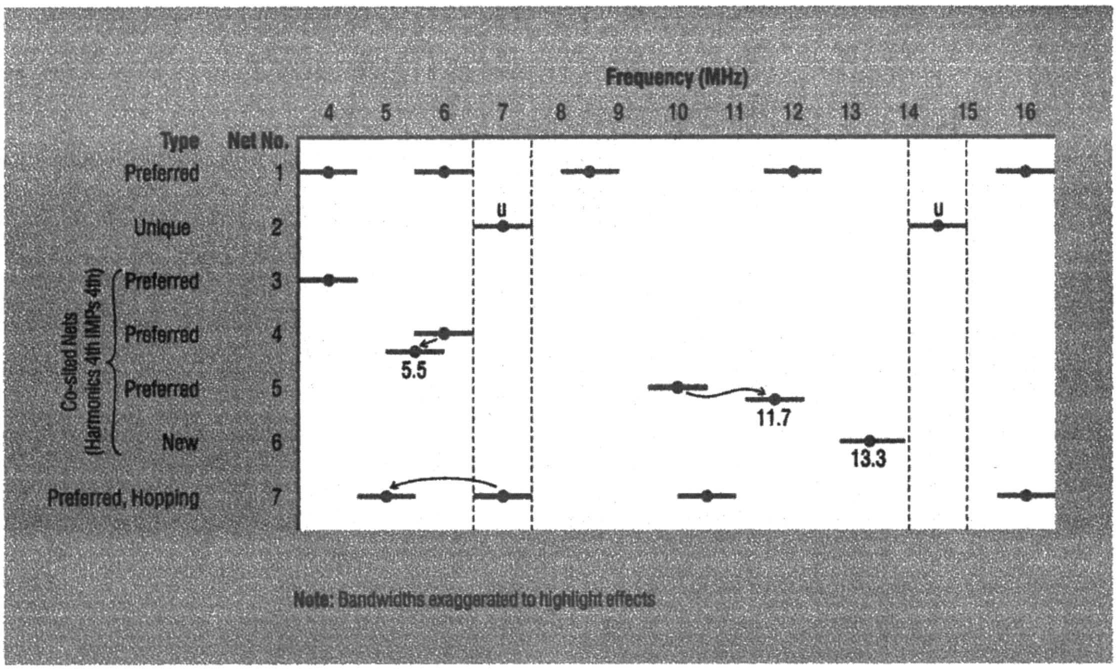


Figure 7. Assignment conflicts.

PAPER TITLE : **Paper 7 - HF Frequency Management: Prediction and Assignment Tools**

AUTHOR : N.S. Wheadon

NAME : D. Jaeger

QUESTION :

Are there some more data available about the differences between the predictions and practical results, especially in the HF-band? (S/Ns).

ANSWER :

The ITU-R methods have been used within these programs because the algorithms have been verified using databases of world-wide signal measurements. Typically the sky-wave algorithm has been shown to have an 10dB standard deviation of the signal prediction.

PAPER TITLE : **Paper 7 - HF Frequency Management: Prediction and Assignment Tools for Large Networks**

AUTHOR : N.S. Wheadon

NAME : K.S. Kho

QUESTION :

Do you need the background noise levels for HF sky-wave prediction?

If yes, how is the interference taken into account, since all of HF channels are (formally) occupied by someone?

ANSWER :

1. The background noise comprises 3 components at HF: Atmospheric noise due to electrical thunderstorms, galactic noise (>18 MHz) and man-made noise due to machinery, power cables etc. Man-made noise dominates in urban, aircraft and ship environments. The programmes take into account all of these sources.

2. The FMT takes into account interference from external sources provided the sources are included in the database. The WinHF-programme does not perform interference conditions. WinHF performs the S/W calculation and it is up to the operator to select the operating frequency using the predictions, the list of assigned frequencies and the observations of on-air interference.

RADARS BASSE FREQUENCE
COMPATIBILITE AVEC LES AUTRES MOYENS ELECTROMAGNETIQUES

J.L. Zolésio
B. Olivier
Thomson-CSF Airsys
Airsys/RD/DTO/RNO
7-9 rue des Mathurins
92221 Bagneux Cedex, France

1. PRESENTATION

L'évolution des menaces militaires conduit à concevoir des systèmes de détection et d'évaluation de la menace dont les capacités doivent sans cesse croître alors que les contraintes économiques deviennent de plus en plus fortes. En effet, pour les systèmes de surface, la menace se caractérise aujourd'hui par les caractéristiques suivantes :

- SER réduites (missiles, avions furtivisés voire furtifs, ...),
- domaines de vol étendus (de la très basse altitude pour les missiles de croisière jusqu'au zénith pour les missiles ARM),
- dynamiques de vol accrues qui permettent d'exploiter des masques naturels du terrain (hélicoptères et missiles de croisière) et/ou des vitesse très élevées,
- des potentialités de contre mesure électroniques accrues avec du brouillage classique mais aussi le masquage de signatures par des moyens passifs (absorbants, formes adaptées) ou actifs.

Les solutions radar continuent à rester les mieux adaptées compte tenu des exigences de portée et de fonctionnement tous temps, le recours aux bandes de fréquence inférieure à 1GHz constituant une solution particulièrement intéressante pour s'adapter aux évolutions de la menace. En effet dans ces bandes de fréquence on bénéficie des avantages suivants :

- SER des petites cibles (missiles) renforcées (phénomène de résonance),
- Effet des masquages sur le bilan radar réduits (détection à TBA),
- Effet de furtivité des cibles réduits,
- Bilan de liaison possible avec des antennes peu directives (couverture instantanée importante et notamment des sites élevés).

L'exploitation de bandes basses en radar permet de plus de bénéficier d'une technologie duale avec les télécommunications.

Différents démonstrateurs de radars bande basse ont déjà été réalisés et ont permis de vérifier l'intérêt de ces bandes de fréquence en radar. Aussi, il convient à présent de préparer l'introduction de tels radars dans les forces et notamment leur intégration dans les plans de fréquence.

Les forces armées bénéficient de nombreuses bandes de fréquence réservées entre 200 et 500 MHz. Le nombre de canaux affectés aux besoins de radiolocalisation est toutefois limité, les applications principales étant du domaine des radiocommunications. La bande 225-400 Mhz est harmonisée par l'OTAN pour applications radiocommunications (relais de télécommunication tactiques : RITA pour la France ou communications air/air ou air/sol).

Dans ces conditions, les radars militaires devront être capables d'utiliser soit des canaux affectés à la radiolocalisation soit ceux affectés à la radiocommunication en prenant alors soin de ne pas engendrer de gêne pour les radiocommunications (fonctionnant simultanément avec le radar). Cette exigence spécifique de ce type de radar conduit à mettre en oeuvre différentes techniques adaptées notamment en ce qui concerne l'agilité de fréquence et la recherche en temps réel de fréquence à CEM Optimale. Il convient également d'exploiter des formes d'onde et des dispositifs de réception adaptés aux exigences de CEM.

2. SOLUTIONS TECHNIQUES

Pour garantir la cohabitation des radars avec les autres moyens électromagnétiques du champ de bataille il convient par ordre de priorité :

- que le radar ne perturbe pas le fonctionnement des autres moyens électromagnétiques,
- que le radar fonctionne normalement en présence des autres moyens électromagnétiques du champ de bataille.

A cette fin il s'agit tout d'abord, de conserver les canaux militaires dans les bandes 200-500 Mhz, ceux réservés respectivement à la radiolocalisation et aux radiocommunications.

De plus, il faut mettre en oeuvre de façon conjointe des caractéristiques techniques particulières et des stratégies d'occupation du spectre « intelligentes ».

2.1 Caractéristiques techniques particulières.

Il existe trois caractéristiques techniques essentielles:

- le filtrage spatial,
- la sélectivité fréquencielle,
- l'utilisation de formes d'onde particulières.

Le filtrage spatial.

Les radars prévus présentent généralement des antennes directives et dans ces conditions seule la zone géographique fixée par le lobe d'antenne est à considérer pour les besoins de CEM.

Concernant l'impact de l'émission radar sur les systèmes de communication, le diagramme naturel d'antenne permet de réduire de 20 dB environ la puissance rayonnée dans les directions autres que celles visées.

Concernant l'impact d'émissions de radiocommunications sur le radar, on peut mettre en oeuvre des techniques de Formation de Faisceau par le Calcul qui permettent d'augmenter encore le niveau de réjection de directions de l'espace autres que celle visée.

La spécificité des radars basse fréquence réside ici dans

Paper presented at the RTO IST Symposium on "Frequency Assignment, Sharing and Conservation in Systems (Aerospace)", held in Aalborg, Denmark, 5-7 October 1998, and published in RTO MP-13.

le faible nombre de capteur des antennes exploitées. AIRSYS a validé, avec le concours du SPOTI, les performances d'algorithmes de ce type.

La sélectivité fréquencielle.
Dans le principe, il s'agit d'une part de concentrer l'émission radar dans un ou plusieurs canaux en limitant au maximum l'énergie rayonnée en dehors de ce ou ces derniers et d'autre part de réduire la réception radar aux canaux exploités en émission. L'objectif est de juxtaposer au plus court des émissions radio et radar exploitant des canaux adjacents.
A l'émission radar des formes d'impulsion à décroissance spectrale rapides sont utilisées.
A la réception, il s'agit de renforcer la sélectivité des étages de réception. Une optimisation sélectivité/facteur de bruit doit cependant être trouvée pour chaque application radar envisagée.

L'utilisation de formes d'onde particulières
Il s'agit de limiter la puissance crête rayonnée par le radar afin de limiter a distance donnée l'impact de l'onde radar sur les systèmes de communication. Pour cela on a largement recours à la compression d'impulsion, la mise en oeuvre de formes d'ondes imbriquées (impulsion courte et impulsion longue) permettant de garantir simultanément une zone aveugle acceptable.
D'autres solutions consistent à émettre des impulsions à spectre en peigne (exploitation de plusieurs canaux libres non adjacents) bien qu'elle conduisent à une complexité et par conséquent à un coût supérieurs (exploitation d'une large bande de fréquence avec occultation de canaux de radiocommunication).

2.2 Stratégie d'occupation du spectre « intelligente »

Deux principaux modes de fonctionnement doivent être envisagés.

Le premier qui est également le plus naturel consiste à prendre en compte les radars dans un plan de fréquence LOCAL. Une analyse du contexte champ de bataille montre, en effet, que pour un radar sol/air courte ou moyenne portée, il convient de coordonner son émission avec les deux ou trois relais de communication tactiques les plus proches (RITA pour la France), les autres relais plus éloignés pouvant fonctionner sur les mêmes canaux de fréquence que le radar sans gêne mutuelle [1].

Le deuxième mode de fonctionnement peut être qualifié de mode agile. Dans ce cas, le radar adapte son émission à l'environnement instantané en s'appuyant sur une écoute préalable des émissions présentes mesurées par son antenne (voir filtrage spatial). Suite à cette écoute, on sélectionne une bande de fréquence permettant de garantir un écart suffisant avec les différentes émissions présentes. Cet écart prend en compte les exigences de non perturbation des équipements de radiocommunication et de bon fonctionnement du radar.

3. CONCLUSION

Les travaux menés par AIRSYS depuis de nombreuses années avec le concours de l'administration Française ont permis de valider l'intérêt de radars basse fréquence pour les systèmes sols du champ de bataille.

Différents travaux ont été menés pour préparer l'intégration de ce type de radar dans les systèmes de défense actuels ou futurs. Ces travaux ont permis de définir un ensemble de techniques à mettre en oeuvre dans ces radars afin de garantir simultanément leur bon fonctionnement et celui des systèmes de communication qui occupent partiellement les bandes V/UHF. Par ailleurs des contraintes acceptables d'établissement d'un plan de fréquence local en temps de paix ou d'utilisation de formes d'onde adaptées et d'occupation spectrale à CEM optimale en temps de crise garantissent la CEM de ce type de radar et des moyens de radiocommunications.

Plusieurs démonstrateurs existent déjà, ils ont été exploités dans de nombreuses expérimentations y compris en zone urbaine et à proximité de bases aériennes ou d'aéroports. Aucun problème particulier de perturbation d'équipements de radiocommunication n'a été noté.

A l'issue de ces travaux il apparaît que les radars VHF ne nécessitent que quelques créneaux de fréquence dans les bandes V/UHF et qu'ils peuvent aisément se placer de manière automatique sur des créneaux disponibles. Cette fonctionnalité est garantie par la possibilité de réaliser des systèmes d'émission réception large en V/UHF.

AIRSYS a développé des briques de base de radars VHF permettant de réaliser des démonstrateur à moindre coût et avec des délais réduits pour préparer le développement de radars opérationnels exploitant les basses fréquences. Les sous-ensembles d'émission et de réception prennent en compte les solutions techniques décrites ci-dessus pour répondre aux exigences de CEM.
Ces radars fonctionneront sur des canaux de fréquence alloués aux forces armées pour applications de radiocommunication et de radiolocalisation qui doivent donc être au minimum maintenus, voire étendus.

---
[1] résultat obtenus grace aux mesures techniques citées précédemment

PAPER TITLE : **Paper 8 - Compatibility of Low Frequency Radar with Other Electromagnetic Assets**

AUTHOR : J-L. Zolesio, B. Olivier

NAME : M. Elliott NATO HQ C3 Staff

QUESTION :

From your briefing I take it that you realise that only the 225-400 MHz band provides a potential frequency resource for this type of radar. Would you prefer frequencies towards the lower or upper edges of this band?

2.Could you say something about the possible planning for the introduction of this type of radar into service?

ANSWER :

Le choix de la bande doit être considéré d'application à application. La partie basse est mieux adaptée pour la propagation en présence de micro-relief alors que la partie haute conduit à des antennes de plus petites dimensions.

Les modules disponibles à AIRSYS permettent de réaliser des expériences en démonstrateur avec des détails de 2 cm environ.

NAME : M. Elliott NATO HQ C3 Staff

QUESTION :

What size antenna are we discussing, and how can the antenna gains described be obtained in this Frequency range?

ANSWER :

Les antennes sont du type "longue bande" (brevet Thomson). Les dimensions des antennes sont typiquement de l'ordre que quelques mètres; celles-ci sont définies par les besoins de taclicité.

PAPER TITLE : **Paper 8 - Compatibility of Low Frequency Radar with Other Electromagnetic Assets**

AUTHOR : J-L. Zolesio, B. Olivier

NAME : K.S. Kho

QUESTION :

1. In your presentation you mentioned:-

   a. It has a module to control EMC with Comms in the band
   b. You need several channels in the R/R sub-band
   c. You have also FH Mode for ECCM

   For FH, how many channels are sequenced and whether R/R sub-band will be sufficient? Could you explain the relationship of the above elements?

2. If you do not apply ECCM (FH), would your radar not be very vulnerable against simple UHF jammers?

ANSWER :

L'aspect contre mesure ne fait pas partie de l'exposé. On cherche à limiter le nombre de canaux alloués et par conséquent l'évasion de fréquence ne doit être utilisée qu'avec précaution, surtout pour permettre au radar de se placer où on l'autorise et ceci sur une large plage de fréquences.

Optimising Radio Network Design

Simon J. Chapman,
Steve Hurley
Rupert Kapp-Rawnsley

Department of Computer Science
University of Wales Cardiff
PO Box 916, Cardiff CF2 3XF
United Kingdom.
Email: {simon,steve,scmrkr}@cs.cf.ac.uk

1 Introduction

Designing radio networks that utilise their allocated frequencies effectively and efficiently is a difficult problem. If a radio network is poorly designed then spectrum will be wasted and/or the quality of service will be degraded even if a good frequency assignment algorithm is used. To produce a well designed network the designer needs to take into account several competing factors. For example, cost may be reduced by having a few omni-directional antennas operating at full power, this may produce good area coverage and have a small amount of overlap between areas (and hence low interference). However, such a network may not be able to satisfy the traffic demands within the area assigned to each antenna i.e. its *cell*. To try and overcome this problem more antennas are required (perhaps using directional antennas at the same site or additional antennas at different sites). However this increases the cost, the potential for interference, and increases the difficulty of finding a good frequency assignment. For example, if the network design is used to generate channel separation constraints between pairs of transceivers then the required separations could have higher values on a poorly designed network relative to a well designed network. Consequently, frequency assignment algorithms e.g. [5, 4] will find assignments which either use a larger range of frequencies than may be necessary (for minimum span assignment) or have a higher number of constraint violations in fixed spectrum problems.

The radio network optimisation problem involves designing a radio network using an optimisation algorithm that takes into account competing factors. For example, the final network can be optimised for cost, interference, handover and traffic demands. Other constraints can be included as necessary. Some related work using models of varying complexity appear in [2, 7, 8, 1, 6, 3].

2 The Generic Model

Described below is a network model. The network is defined within a working area $\mathcal{P}$. Any values that can be varied are framed at their first occurrence.

Any point on $\mathcal{P}$ is defined by coordinates (x, y). For simplicity, points on $\mathcal{P}$ are only defined on a grid. Such data we shall refer to as mesh-data.

Points within $\mathcal{P}$ that have been tested for reception are known as reception test points. These are represented by $\mathcal{R}$, where

$$\mathcal{R} = \{R_1, R_2, \cdots, R_{n_\mathcal{R}}\}.$$

$n_\mathcal{R}$ is the total number of reception test points. We represent a typical reception test point by R_i where $1 \leq i \leq n_\mathcal{R}$

Mesh-data is provided for the following:

- relative elevation,
- propagation loss estimates,
- service, and
- traffic.

Paper presented at the RTO IST Symposium on "Frequency Assignment, Sharing and Conservation in Systems (Aerospace)", held in Aalborg, Denmark, 5-7 October 1998, and published in RTO MP-13.

Other data, which we shall refer to as engineering-data, is provided for the following:

- candidate site positions,
- candidate site legacy status,
- antenna types defined by horizontal and vertical diagrams as well as transmission gain and loss.

Both the mesh- and engineering-data are described in more detail below.

2.1 A Network.

Consider a number of candidate sites

$$\mathcal{L} = \{L_1, L_2, \cdots, L_{n_{\text{candidate}}}\}.$$

$n_{\text{candidate}}$ is the total number of candidate sites.

A network, $\mathcal{N}$, is composed of sites chosen from the list of candidate sites $\mathcal{L}$ ($\mathcal{N} \subseteq \mathcal{L}$). To define a network we define an "on gene", $\boxed{\mathcal{Y}}$, where

$$\mathcal{Y} = \{y_1, y_2, \cdots, y_{n_{\text{candidate}}}\},$$

where, for the l^{th} candidate site in $\mathcal{L}$,

$$y_l = \begin{cases} 0, & \text{if site } L_l \text{ is "off", } i.e. \text{ it is not used,} \\ 1, & \text{if site } L_l \text{ is operational, } i.e. \text{ it is used (contains at least one station).} \end{cases}$$

Therefore a network is defined as follows,

$$\begin{aligned} \mathcal{N} &= \{L_l : y_l = 1, L_l \in \mathcal{L}, 1 \leq l \leq n_{\text{candidate}}\} \\ &= \{N_1, N_2, \cdots, N_{n_{\text{operational}}}\} \\ &= \{N_m : 1 \leq m \leq n_{\text{operational}}\}. \end{aligned}$$

N_m represents an operational site and $n_{\text{operational}} \leq n_{\text{candidate}}$.

In cases where an existing network is to be altered to cope with new conditions we have to take into account the presence of legacy, or currently operational, sites. Therefore, in addition to the "on gene" we have to provide a "legacy gene", $\mathcal{Z}$, where, for the l^{th} candidate site in $\mathcal{L}$,

$$\mathcal{Z} = \{z_1, z_2, \cdots, z_{n_{\text{candidate}}}\},$$

where,

$$z_l = \begin{cases} 0, & \text{if site } L_l \text{ is a new candidate, } i.e. \text{ it is not part of an existing network,} \\ 1, & \text{if site } L_l \text{ is "legacy", } i.e. \text{ it is part of an existing network.} \end{cases}$$

A network provides a service based upon criteria defined by the network operator. The nature of the expected service for a network defined on $\mathcal{P}$ is given by $\mathcal{S}$. The set $\mathcal{S}$ is defined as the set of points:

$$\mathcal{S} = \{S_1, S_2, \cdots, S_{n_\mathcal{S}}\}.$$

$n_\mathcal{S}$ is the total number of service test points. We represent the service at reception test point R_i by S_i.

Also during operation the network must handle the traffic in an efficient way. The expected traffic demand for a network defined on $\mathcal{P}$ is given by $\mathcal{T}$. The set $\mathcal{T}$ is defined as the set of points:

$$\mathcal{T} = \{e_1, e_2, \cdots, e_{n_\mathcal{T}}\},$$

where each point in $\mathcal{T}$ is measured in Erlang. $n_\mathcal{T}$ is the total number of traffic test points. We represent the traffic, measured in Erlang, at reception test point R_i by e_i.

We have the following inclusion:

$$\mathcal{T} \subseteq \mathcal{S} \subseteq \mathcal{R},$$

and

$$n_\mathcal{T} \leq n_\mathcal{S} \leq n_\mathcal{R}.$$

2.2 A Site.

The monetary cost of a site is represented by

$$\mathcal{F} = \{f_1, f_2, \cdots, f_{n_{\text{candidate}}}\}.$$

Each operational site consists of one or more base stations (BS). The model represents a base station by

$$\gamma_{mk_m}, \text{ where } k_m \in \{1, 2, \cdots, \boxed{n_m^{\text{BS}}}\}$$

where γ_{mk_m} is the k_m^{th} base station of operational site N_m, and where n_m^{BS} is the number of base stations at site N_m. For an operational site $n_m^{\text{BS}} \geq 1$.

A candidate site, L_l, has position on $\mathcal{P}$ represented by

$$\mathfrak{r}_l = (x_l, y_l, h_l)$$

where x_l and y_l are the x- and y-coordinates (in Lambert II coordinates) of candidate site L_l, and h_l its height above sea-level. Instead of the height we use an angular elevation mesh.

The angle of elevation to each point defined on $\mathcal{P}$ relative to each site is given by $\mathcal{E}$. Where,

$$\mathcal{E} = \{\alpha_1, \alpha_2, \cdots, \alpha_{n_{\text{candidate}}}\}.$$

α_l is the angle of elevation to each mesh-point defined on $\mathcal{P}$ relative to site L_l.

The propagation loss to each mesh-point defined on $\mathcal{P}$ relative to each site is given by $\mathcal{Q}$. Where,

$$\mathcal{Q} = \{Q_1, Q_2, \cdots, Q_{n_{\text{candidate}}}\}.$$

Q_l is an estimate of the propagation losses to each mesh-point defined on $\mathcal{P}$ relative to site L_l.

2.3 A Base Station.

The model considers a set of base stations, $\mathcal{B}$, such that:

$$\mathcal{B} = \{B_1, B_2, \cdots, B_{n_\mathcal{B}}\},$$

where $n_\mathcal{B}$ is the total number of base stations.

The k_m^{th} operational base station of site m is represented by γ_{mk_m}. Hence, we have

$$\gamma_{mk_m} = B_j \text{ if } j = \sum_{m'=1}^{m} n_{m'}^{\text{BS}} + k_m.$$

A given base station, B_j where $1 \leq j \leq n_\mathcal{B}$, situated at site N_m has a number of operational parameters that we are free to change.

- $\boxed{B_j^{P_s}}$ is the transmitting power of B_j.
- $\boxed{B_j^{\text{AT}}}$ is the antenna type of B_j.
- $\boxed{B_j^{\beta}}$ is the tilt of the antenna of B_j.
- $\boxed{B_j^{\delta}}$ is the azimuth of the antenna of B_j.
- $\boxed{B_j^{\text{TRX}}}$ is the number of TRX devices used by B_j.

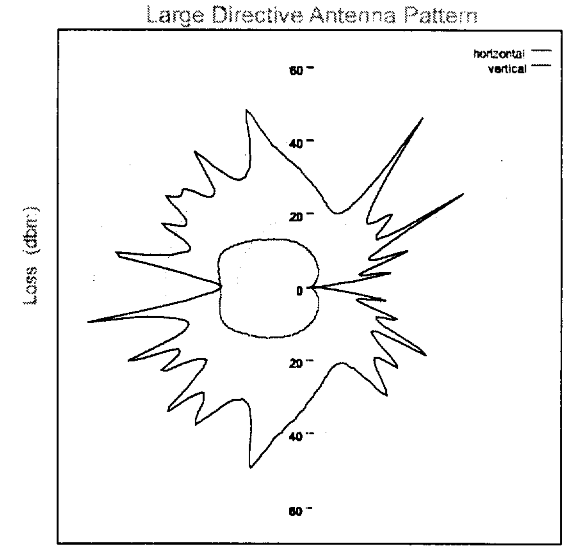
2.4 Antennæ.

A set of antennæ, $\mathcal{A}$, is defined such that:

$$\mathcal{A} = \{A_1, A_2, \cdots, A_{n_\mathcal{A}}\},$$

where $n_\mathcal{A}$ is the total number of antenna types.

Each base station, B_j, has one antenna. An antenna can be omni-directive, narrow panel directive or large panel directive. The "directiveness" of an antenna is characterised by a radiant loss diagram. To fully describe a directive antenna's radiant losses as a function of angle requires two diagrams - horizontal and vertical (see Figure 1 as an example).



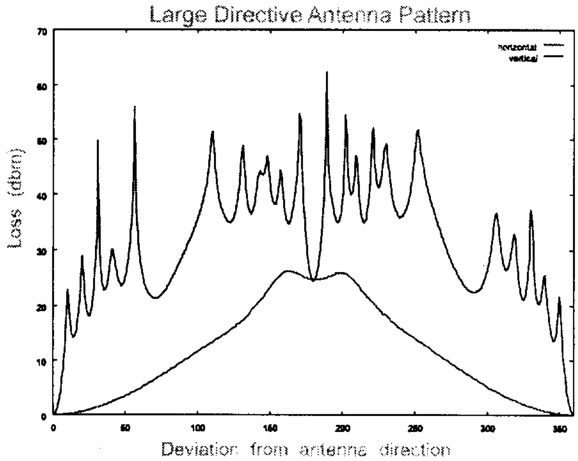


Figure 1: Radiant losses for a large panel directive antenna

Therefore, for a given antenna type we have radient loss characteristics

$$D_{\text{horizontal}},$$

and

$$D_{\text{vertical}}.$$

Each antenna has an additional fixed transmission gain and loss represented, respectively, as

$$G_\mathcal{A} \text{ and } \Lambda_\mathcal{A}.$$

2.5 Mobile.

A set of mobile stations, $\mathcal{M}$, is defined such that:

$$\mathcal{M} = \{M_1, M_2, \cdots, M_{n_\mathcal{M}}\},$$

where $n_\mathcal{M}$ is the total number of mobile stations. Each mobile station has an associated reception gain and loss represented, respectively, by

$$g_\mathcal{M} \text{ and } \lambda_\mathcal{M}.$$

The mobile type at a given reception test point is given in the set of service test points, $\mathcal{S}$. That is the mobile type is a function of S_i. Therefore, we represent a mobile at a given service test point S_i by M_i (i.e. the mobile to be serviced at S_i).

2.6 Definitions.

2.6.1 Downlink Field Strength.

For a given base station, B_j, situated at site N_m the downlink field strength at a reception point, R_i, is given by

$$\begin{aligned} F_j(R_i) &= B_j^{P_s} + G_\mathcal{A}(B_j^{\text{AT}}) - \Lambda_\mathcal{A}(B_j^{\text{AT}}) \\ &- Q_m(R_i) \\ &- D_{\text{horizontal}}(B_j^{\text{AT}}, B_j^\delta, R_i) \\ &- D_{\text{vertical}}(B_j^{\text{AT}}, B_j^\beta, \alpha_j(R_i)) \\ &+ g_\mathcal{M}(M_i) - \lambda_\mathcal{M}(M_i). \end{aligned}$$

2.6.2 Cell.

A cell, C_j, is defined for base station B_j as the set of points such that:

$$C_j = \{R_i : F_j(R_i) \geq S_i\}.$$

2.6.3 Kernel.

The kernel, ker_u, operates on the cell, C_j, of base station B_j and is defined as the subset of points, R_i, in C_j such that the field strength at R_i recieved from base station B_j is at least u above the field strength at each R_i received from all other base stations, $B_{j'}$ $j' \neq j$. That is,

$$\begin{aligned} \text{ker}_u C_j = \; & \{R_i : F_j(R_i) \geq S_i \text{ and } \forall j' \neq j \\ & F_j(R_i) - F_{j'}(R_i) \geq u\}. \end{aligned}$$

2.6.4 Handover.

The handover, hand_u, operates on the cell, C_j, of the base station B_j and is defined as the subset of points, R_i, in C_j such that the field strength at R_i recieved from base station B_j is at least within u of the field strength at each R_i received from a neighbour base station $B_{j'}$. That is,

$$\begin{aligned} \text{hand}_u C_j = \; & \{R_i : R_i \in C_j \setminus \text{ker}_u C_j \text{ and } \exists\, j' \neq j \\ & \text{s.t. } |F_j(R_i) - F_{j'}(R_i)| \leq u\}. \end{aligned}$$

For a given cell the handover may be satisfied by more than one base station. The target number of neighbour base stations that overlap a cell's handover region is given by

$$h_j.$$

2.6.5 Connectivity.

A set of points, $\mathcal{X}$, is fully connected if there exists some path joining any two points within $\mathcal{X}$. We represent this by $\mathcal{X}^c = 1$.

If a fully connected set does not exist but that set when split into two disjoint subsets forms two fully connected subsets then we represent this by $\mathcal{X}^c = 2$ and describe each subset as a component of $\mathcal{X}$.

Therefore, $\mathcal{X}^c = n$ means that the set $\mathcal{X}$ is composed of n connected components.

In our case we are interested in the connectivity of the entire network where a service is defined, $\mathcal{S}^c$, and the connectivity of individual cells, C_j^c.

2.7 Constraints and Objectives.

2.7.1 Existing Network.

If existing sites cannot be decommissioned then we have the following constraint

$$\text{if } z_l = 1 \text{ then } y_l = 1 \text{ always.}$$

2.7.2 Number of Sites.

The number of sites are constrained as follows

$$\min \sum_{l=1}^{n_{\text{candidate}}} y_l$$

2.7.3 Cost.

The total monetary cost is constrained as follows

$$\min \sum_{l=1}^{n_{\text{candidate}}} f_l y_l$$

2.7.4 Cover.

For all service test points, R_i, we require

$$\sum_{j=1}^{n_B} x_{ij} = \begin{cases} 1, \text{ if } R_i \in \text{ker}_u C_j, \\ h_j, \text{ if } R_i \in \text{hand}_u C_j \end{cases}$$

such that

$$\text{hand}_u C_j \neq \emptyset,$$

and

$$\text{ker}_u C_j \neq \emptyset,$$

where

$$x_{ij} = \begin{cases} 1, \text{ if } R_i \in C_j, \\ 0, \text{ else.} \end{cases}$$

2.7.5 Connectivity.

We require that all service test points defined on $\mathcal{P}$ be connected. That is,

$$\mathcal{S}^c = 1.$$

We also require that a cell be fully connected. That is, the cell has only one component:

$$C_j^c = 1.$$

2.7.6 Traffic.

We maximise the traffic across a cell. That is, for each cell, C_j,

$$\max \frac{1}{(B_j^{\text{TRX}})_{\text{Erlang}}} \sum_{R_i \in C_j} e_i.$$

$(B_j^{\text{TRX}})_{\text{Erlang}}$ is the capacity, measured in Erlang, of base station B_j produced by B_j^{TRX} TRX devices.

3 Network Initialisation

The initialisation process differs based upon the type of network design that is required. We distinguish between two, *greenfield* and *expansion.*

Greenfield regions contain no initial or starting network. The starting network for such a region is initialised as follows:

- The theoretical minimum number of antennae, N_{min}, that are needed to satisfy the total network demands of the proposed network, is calculated as

$$N_{min} = \left\lceil \frac{\sum_i^{n_T} e_i}{\mathcal{T}_{max}} \right\rceil.$$

  Where $\mathcal{T}_{max}$ is the maximum traffic that a single antenna can service.

- The starting number of antennae for the proposed network, N_{start}, is calculated as

$$N_{start} = \beta N_{min}.$$

  β is a control parameter in the range $0 \leq \beta \leq \beta_{max}$, where

$$\beta_{max} = \frac{n_{candidate} N_{antenna}^{max}}{N_{min}}$$

  $n_{candidate}$ is the total number of potential sites, and $N_{antenna}^{max}$ is the maximum number of antennae allowed at one site. If β is low the network is built via an "additive" process; while if β is large the network is built via a "subtractive" process.

- N_{start} omni-directional antennae are then positioned at randomly chosen sites.

In contrast to greenfield regions an expansion region contains an existing, legacy, network. The design requirement is to extend or improve this network in order to satisfy new constraints and objectives. The initialisation process simply involves using the legacy network as the starting network.

4 Network Design

Once a starting network has been chosen the design process can begin. During the design a number of "trial" networks are generated from previous networks by investigating a network's neighbourhood, which could be generated randomly or involve more intelligence. Then based upon the *cost*, C, of this trial network, a decision, controlled by a simulated annealing algorithm, is made as to whether an improvement has been obtained. If an improvement occurred the trial network is kept. If no improvement occurred (modulo the simulated annealing algorithm) then it is discarded and the previous best network retained.

The design process is currently identical for both greenfield and expansion regions.

5 Cost Function

The current measure of performance for a trial network is based on a weighted, additive cost of several components. Our multi-objective function takes into account several factors as listed below. Others can be added as necessary. The goal is to minimise the overall cost, C, which is given by

$$C = \sum_i \omega_i C_i$$

Where $\omega_i \geq 0$ are weighting factors.

5.1 Site cost (global)

The monetary cost of all the sites in a network is given by

$$C_1 = \frac{\sum_i \delta_1^i f_i}{\sum_i f_i}$$

Where f_i is the cost of deploying site i and $\delta_1^i = 1$ if site i is used and 0 otherwise. We use

$$f_i = \begin{cases} 1, & \text{if there is no change on site } i \text{ with respect to the initial network;} \\ 2, & \text{if there is any change on site } i \text{ with respect to the initial network;} \\ 5, & \text{if site } i \text{ is newly installed with respect to the initial network;} \\ 7, & \text{if site } i \text{ is removed with respect to the initial network.} \end{cases}$$

If a greenfield network design is required then f_i is automatically set to 5.

5.2 Coverage (global)

The coverage cost is given by

$$C_2 = \frac{n_{hole}^{STP}}{n_{total}^{STP}}$$

Where n_{hole}^{STP} is the total number of STP that are not assigned to a cell and n_{total}^{STP} is the total number of STP.

5.3 Traffic coverage (global)

We are currently experimenting with two global, traffic-based cost functions. The first considers the traffic coverage across the set of TTP.

$$C_{3a} = \frac{n_{hole}^{TTP}}{n_{total}^{TTP}}$$

Where n_{hole}^{TTP} is the total number of TTP that are not covered by any cell that has total traffic $\leq \mathcal{T}_{max}$, where $\mathcal{T}_{max}$ is the maximum traffic that a cell can service, and n_{total}^{TTP} is the total number of TTP.

The second considers the actual traffic coverage across a cell.

$$C_{3b} = \frac{\sum_i \delta_3^i \mathcal{T}_i}{\sum_i \mathcal{T}_i}$$

Where $\mathcal{T}_i$ is the total traffic enclosed by cell i, and $\delta_3^i = 1$ if $\mathcal{T}_i \leq \mathcal{T}_{max}$, and 0 otherwise.

5.4 Handover (global)

The handover objective requires that each cell, i, have h_i^{target} neighbour cells capable of providing handover.

$$C_5 = \sum_i \mid h_i^{target} - h_i \mid$$

Where h_i is the actual number of handover cells for cell i.

5.5 Overlap (global)

The *overlap* objective requires that each cell, i, have zero neighbour cells, excluding handover cells, that

Total traffic	3211.0
Total number of sites	250
Minimum number of stations	75
Number STP	29954
Number TTP	4967

Table 1: Network details

Site	ω_1	0.1
Coverage	ω_2	1.0
Traffic (TTP)	ω_{3a}	0.0
Traffic (cell)	ω_{3b}	2.0
Handover	ω_5	0.1
Overlap	ω_6	0.5

Table 2: Cost function weights

can service at least one STP within cell i. We consider overlap to be a measure of the interference of a network.

The cost function for overlap is:

$$C_6 = \sum_i \delta_6^i(n_i^{overlap} - h_i) + (1-\delta_6^i)(n_i^{overlap} - h_i^{target})$$

Where $n_i^{overlap}$ is the total number of neighbour cells that can service at least one STP within the region defined by cell i. h_i^{target} is the required number of handover cells, h_i is the actual number of handover cells for cell i and where $\delta_6^i = 1$ if $h_i \leq h_i^{target}$ and 0 otherwise.

6 Preliminary Results

Preliminary results are presented for a network which is detailed in Table 1. The minimum number of stations, N_{min}, is given by

$$N_{min} = \left\lceil \frac{\sum_i^{n_T} e_i}{\mathcal{T}_{max}} \right\rceil .$$

Where $\mathcal{T}_{max}$ is the maximum traffic that a single antenna can service, currently set at 43 Erlangs.

The weights used to define the cost function are given is Table 2.

		Before	After
number of cells	n_{cells}	0	162
total cost	C	∞	0.6449
coverage	$\omega_2 C_2$	1.0000	0.0584
cellular traffic	$\omega_{3b} C_{3b}$	∞	0.1503
handover	$\omega_5 C_5$	0.0000	0.0142
overlap	$\omega_6 C_6$	0.0000	0.2930

Table 3: Greenfield design results

		Before	After
number of cells	n_{cells}	180	173
total cost	C	0.6269	0.4358
coverage	$\omega_2 C_2$	0.0654	0.0511
cellular traffic	$\omega_{3b} C_{3b}$	0.1440	0.0363
handover	$\omega_5 C_5$	0.0102	0.0110
overlap	$\omega_6 C_6$	0.3330	0.2890

Table 4: Expansion design results

6.1 Greenfield Design

To initialise the network β is set to zero i.e. the network will be built in an additive manner. Figure 2 shows the progress of the design in relation to the various cost function components. The total cost decreases along with most of the components. As expected the overlap increases since

- the starting network is empty i.e. overlap free, and
- the overlap cost is weighted relatively low.

Table 3 gives the values of the cost function components for the initial and final network design.

6.2 Expansion Design

Here we use an existing legacy network as the starting network for the design process. Figure 3 shows the progress of the algorithm in relation to the various cost function components. In contrast to the greenfield network the overall cost decreases much more slowly. Table 4 gives the values of the cost function components for the initial and final network design, from which we observe that the final design requires less cells, improves area and traffic coverage, improves overlap but generates slightly worse handover regions.

Acknowledgement

This work is funded under EU ESPRIT Project 23243.

References

[1] P. Calegarie, F. Guidec, P. Kuonen, B. Chamaret, S. Udeba, S. Josselin, and D. Wagner, *Radio network planning with combinatorial algorithms*, ACTS mobil communications, (1996), pp. 707–713.

[2] S. J. Fortune, D. M. Gay, B. W. Kernighan, O. Landron, R. A. Valenzuela, and M. H. Wright, *Wise design of indoor wireless systems*, IEEE Computational Science and Engineering, 2 (1995), pp. 58–68.

[3] T. Fruhwirth, J. Molwitz, and P. Bisset, *Planning cordless business communication systems*, IEEE Expert Magazine, (1996), pp. 662–673.

[4] S. Hurley and D. Smith, *Meta-heuristics and channel assignment*, in Methods and algorithms for channel assignment, R. Leese, ed., Oxford University Press, to appear 1998.

[5] S. Hurley, D. Smith, and S. Thiel, FASoft*: A system for discrete channel frequency assignment*, Radio Science, 32 (1997), pp. 1921–1939.

[6] J. McGeehan and H. Anderson, *Optimizing microcell base station locations using simulated annealing techniques*, in 44th IEEE Vehicular Technology Conference, IEEE, 1994, pp. 858–862.

[7] H. D. Sherali, C. M. Pendyala, and T. S. Rappaport, *Optimal location of transmitters for micro-cellular radio communication system design*, IEEE Journal on Selected Areas in Communications, 14 (1996), pp. 662–673.

[8] K. Tutschuka, N. Gerlich, and P. Tran-Gia, *An integrated cellular network planning tool*, in 47th IEEE Vehicular Technology Conference, IEEE, 1997, pp. 765–769.

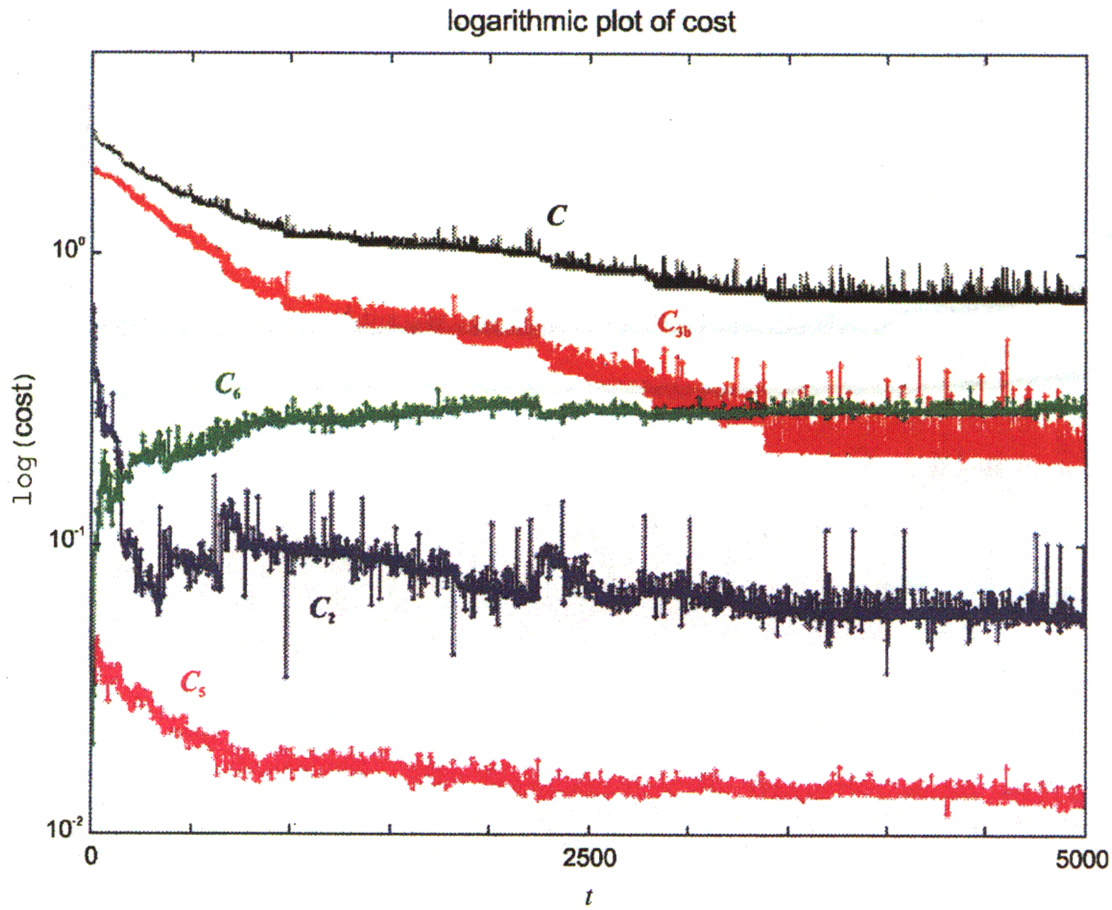


Figure 2. Variation of cost components over 5000 iterations for the greenfield design.

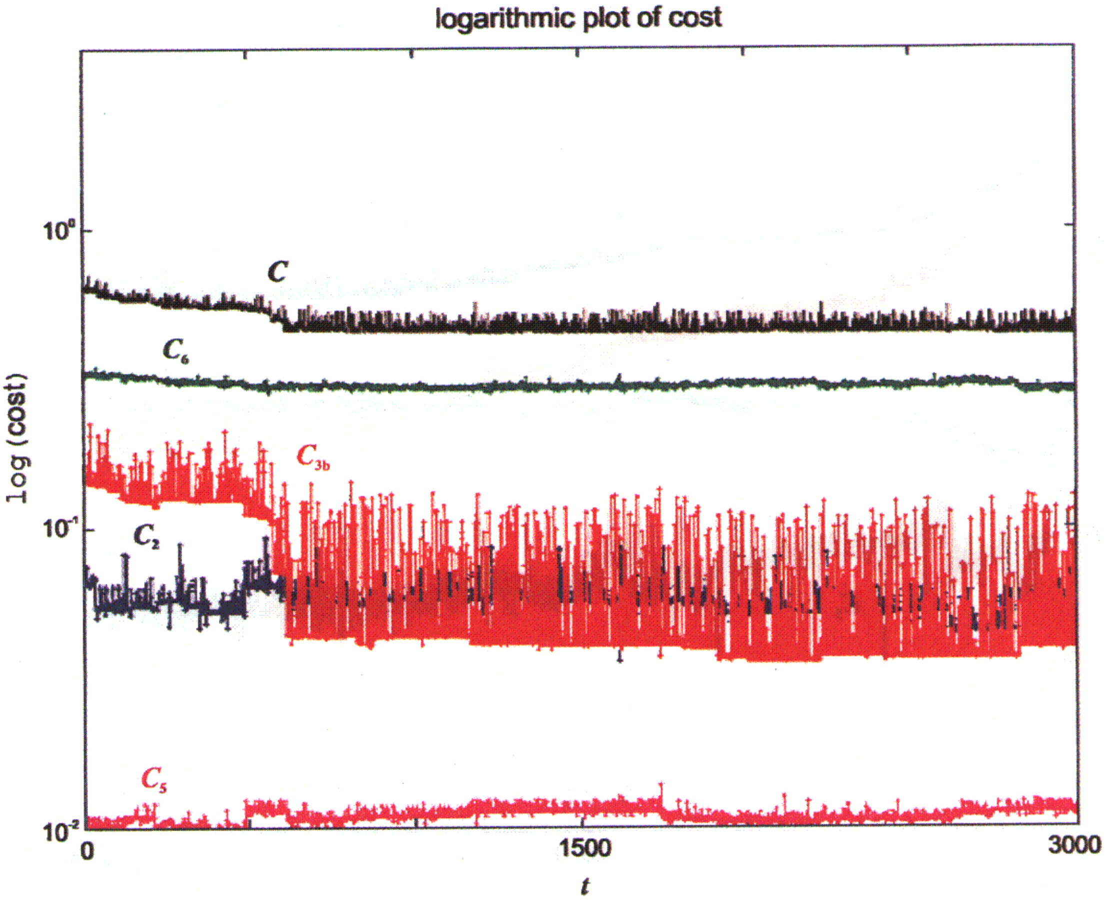


Figure 3. Variation of cost components over 3000 iterations for the expansion design.

PAPER TITLE : **Paper 9 – Optimising Radio Network Design**

AUTHOR : S.J. Chapman et al

NAME : I. White

QUESTION :

1. Why don't your results quantify overlap figures?

2. Where do Meta heuristics come from and how are they used? What can you say about this course and run time?

ANSWER :

1. Simply put, we have not yet done enough work in that direction. The current methods for interference evaluation are speculative. We assume they will be sympathetic to the FAP and hope to quantify this in the future by the application of up to date interference evaluation methods.

2. The meta-heuristics are standard 'themes' for search algorithm design. They are applied and analysed for weaknesses, adjustments are then made to improve their performance.

   In general, 'hand-crafted' heuristics will perform better than meta heuristics because they are fitted closely to the problem structure. However, heuristics which take account of problem specific information will be programmer intensive to employ. Meta-heuristics are computationally intensive (a cheap commodity) and may be employed 'off-the-shelf', to a wide variety of problems.

The Impact of Protection Criteria and Assignment Order on Military Air Ground Air Frequencies

K.S. Kho, M. Elliot
NATO HQ C3 Staff
Frequency Management Branch
1110 Brussels, Belgium

1 Introduction

Tactical air communications use the UHF 225 – 400 MHz military band in NATO Europe. Overall management of the military use of the band is performed by the NATO Frequency Management Sub-Committee (FMSC). The band is divided into allotments for each type of system e.g. Radio Relay (R/R), [Instrument Landing System (ILS),] UHF Satellite and Air – Ground – Air (A/G/A), based on an Allotment Plan agreed by the FMSC. The management of the A/G/A assignments is then performed centrally by the Frequency Management Branch (FMB) of the NATO HQ C3 Staff at NATO HQ, using a software tool called NATO UHF Frequency Assignment System (NUFAS 2).

This paper first describes the assignment system of NUFAS 2 and then focuses on the results of an investigation into the impact of assignment order on the results of a bulk assignment process.

Paragraphs 2 to 7 are background information for your convenience.

2 Individual and Batch versus Bulk Assignments

NUFAS 2 has two basic features for assigning frequencies. They are as follows:

a. Bulk Assignments
b. Individual and Batch Assignments

The Bulk Assignment feature is used by the FMB staff in the case where a high number of assignments should be carried out as a part of re-organisation of the UHF band. In this case, an algorithm to order the assignment requests is provided. The degree of constraint of assignments in fulfilling assignments rules, varies a lot. In order to maximise the chance of success in calculating frequencies, the more constraint assignments should be handled before the less constraint ones.

The FMB staff uses the Individual and Batch Assignment feature for day to day assignments to meet requests for air/ground/air frequencies.

3 Type of Assignments

The various types of A/G/A assignments are described below:

- Simple assignments

Simple assignments are assignments for A/G/A communications within a single service volume. The service volumes defined by the users should accurately match the operational area of use of the communications.

- Pooled assignments

A pool of assignments is a set of assignments with associated individual service volumes where a common frequency is required for all members of the pool.

- Pre-assignments

Pre-assignments are assignments where the frequencies are already defined in the requests. Both simple and pooled assignments can be pre-assigned.

4 Assignment Rules

The assignment algorithm of NUFAS 2 system considers the following three aspects:

a) Resource Rules

The availability of frequencies for any request for assignment is investigated by NUFAS 2 in accordance with rules based on the FMSC Allotment Plan for the UHF band. With these rules, the FMSC decisions regarding usage of frequencies, protection to restricted / sensitive frequencies etc. can be realised.

Paper presented at the RTO IST Symposium on "Frequency Assignment, Sharing and Conservation in Systems (Aerospace)", held in Aalborg, Denmark, 5-7 October 1998, and published in RTO MP-13.

b) Cosite Protection

Cosite Protection rules are used to avoid mutual interference between assignments of frequencies located at a single site. The user can define minimal frequency separation, harmonic protection from VHF frequencies on the site and also Intermodulation (IM) protection from 2 signal 2nd order up to 3 signal 5th order IM products.

The cosite constraints concerning the level of IM product and frequency separation can be relaxed up to two times resulting in up to three qualities of assigned frequencies.

c) Intersite Protection

The intersite algorithm calculates the channel separation required between two assignments that are not cosited. The module can determine whether no channel separation, or 1 channel, 2 channels or 3-channel separation is required for the pair of assignments under consideration. The number of channels considered is a program constant that can be set to higher values if required, but with a consequential impact on processing time.

5 Bulk Assignment

When the UHF band is re-organised, a high number of A/G/A frequencies, if not all, should be re-assigned. Therefore the algorithm to perform the task is different from the algorithm used in the case of individual and batch assignments. In bulk assignment, a matrix is created by NUFAS 2 which is used to order the assignment requests in assignment difficulty. The matrix will contain the minimal frequency separation required between each request and each back-ground assignment or request. The assignment order is derived based on the minimal frequency separation values and weighting factors.

6 Matrix and Assignment Order

In the case of bulk assignments, the process of assigning frequencies will only be started after creating this "Bulk Assignment Matrix" and ordering the assignment requests. In principle, the matrix contains the following information:

a. Channel separation required between two assignments
b. The measure of difficulty of the assignment (the "Score"), based on the channel separation values and the weighting factors.

An illustration of a matrix is given below: (Note that the matrix is symmetrical)

XXX	Ass1	Ass2	Req1	Req2	Req3	Req4	Req5	Req6	Req7	Score
Ass1	XXX		4	3	1	2	5	3	2	63
Ass2		XXX	3	4	2	0	3	1	4	28
Req1	4	3	XXX	1	4	4	0	5	3	71
Req2	3	4	1	XXX	2	3	4	0	1	29
Req3	1	2	4	2	XXX	1	2	4	0	25
Req4	2	0	4	3	1	XXX	0	1	4	24
Req5	5	3	0	4	2	0	XXX	0	5	93
Req6	3	1	5	0	4	1	0	XXX	4	60
Req7	2	4	3	1	0	4	5	4	XXX	69

Legend: Ass1 = an existing assignment
Req1 = Request for an assignment
Score = Score using current weighting factors

Pooled assignments are considered as a group. Each member of a pool has its matrix value set to the highest value for the pool members. This process is called the

"Poolification" process. If for example, Req1, Req3 and Req7 are members of a pool, then, after poolification, the score values for Req1, Req3 and Req7 will be then 102. The rows concerning Req1, Req2 and Req3 becomes as follows:

XXX	Ass1	Ass2	Req1	Req2	Req3	Req4	Req5	Req6	Req7	Score
Req1	4	4	XXX	2	XXX	4	5	5	XXX	102
Req3	4	4	XXX	2	XXX	4	5	5	XXX	102
Req7	4	4	XXX	2	XXX	4	5	5	XXX	102

Therefore Req1, Req3 and Req7 will be assigned first and not after the assignment of Req5.

7 Weighting Factor

To reflect the degree of difficulty in assigning frequencies, a weighting factor is applied to each matrix value. The current weighting factor applied is as follows:

Channel Separation Required (channels)	**Weighting factor**
0	0
1	1
2	3
3	5
4	7
5 (cosite)	39

Table 1: Current weighting factor of NUFAS 2 system

The current weighting factors mentioned in table 1 were selected based on the results of an investigation on frequency assignment algorithms carried out in an earlier study into frequency assignment algorithms[1]. This weighting factor equals the number of channels denied to a request by a given required channel separation.

Once the assignment order is determined, the actual assignment of frequencies to requests takes place. For each request, each frequency in the Resource List is examined in a pre-defined order until an assignable frequency is reached that passes all the assignment rules. This frequency is then assigned to the request. Requests are processed in decreasing order of difficulty until all the requests have been assigned. Some requests may remain un-assigned at the end of the process ("failures").

The results of the bulk assignment depend on the weighting factor used. In principle, the weighting system reflects already most of the aspects regarding the degree of difficulty in finding frequencies. However, the trial results indicate that the tuning range is also significant in this respect.

8 Test Method

The investigation that was carried out considered the impact of changing the weighting factors on the results of the bulk assignment process. For each trial and set of weighting factor, the bulk assignment was run. The results were assessed by considering the failures. In addition, graphs of assigned requests against the resource number were used to assess the efficiency of the ordering process.

[1] Dr. T.A. Lanfear
NEMCA Project 5,
'Graph Theory and Radio Frequency Assignment', dated 1989

9 Scenario

The number of A/G/A records in Master Radio Frequency List (MRFL) database was 13056. The distribution of the sorts of records was as follows:

Number of Background assignments: 0
Number of request: 13056 of which :
Pre_assignments: 3161, 2478 of them were in 349 pools
Normal Requests: 9895, 7294 of them were in 1406 pools

It can be seen that most of the assignments were inter-related in pools.

The trials were caried out with five different sets of weighting factors. The weighting factors used area as follows :

Channel Separartion Required (channels)	No Weighting factor (Trial 1)	Current Weighting factor (Trial 2)	Weighting factor (Trial 3)	No Weighting factor Cosite (Trial 4)	No weighting Factor Intersite (Trial 5)
0	0	0	0	0	0
1	0	1	1	1	0
2	0	3	2	3	0
3	0	5	3	5	0
4	0	7	4	7	0
>4	0	39	19	0	39

10 Results

The results of the trials are as follows :

Results	**Trial 1**	**Trial 2**	**Trial 3**
Failures	1 Normal Request and 65 Pooled Requests failed involving 4 pools with 15 to 25 members per pool (5 frequencies)	174 Pooled Requests failed involving 11 pools with 11 to 25 members per pool (11 frequencies)	No failures

Results	**Trial 4**	**Trial 5**
Failures	80 Pooled Requests failed involving 4 pools with 15 to 25 members per pool (4 frequencies)	1 Normal Request failed with very limited tuning range (1 frequency)

The results in terms of failures, were not consistent. Trial 2, which should be the optimal weighting, was even the worst trial. This is caused by the complicated nature of the military frequency records such as pre-assignments, pooled assignments etc. Clearly, the relatively small number of failures is sensitive to the weighting factors choosen.

It can also be seen in trial 5 that a weighting value that only takes into account cosited assignments still leads to a good result.

The distribution of the frequency usage in trial 3 is given in figures 1 to 3 at pages 5 to 6.

Figure 1 shows the distribution of the assigned frequencies. It can be seen that the distribution is random. The frequency resources are scrambled by blocks of 20 channels in a random manner. The empty sub-bands represent sub-bands alloted to other military systems.

Figure 2 displays the frequency re-use rate. The resources with lower numbers are re-used more than the others. This is logical, since the frequency assignment algorithm is based on the frequency exhaustive method. In the next version of the assignment system, NUFAS 3, the algorithm will be expanded with selection based on other criteria such as the best D/U ratio or the least used frequency etc. This new algorithm may lead to a more random distribution of frequency re-use.

Figure 3 displays the resource number as a function of the ordered assignment. Due to the frequency exhaustive algorithm used, frequency resources with lower position numbers are more assigned than the ones with higher position numbers.

It can be seen in figure 3 that resources with numbers between 6000 and 7000 have relatively few assignments. Therefore, it is expected that it would still possible to accommodate some additional assignment requests. Based on the current occupancy of the UHF band, the frequency assignment system will face serious frequency resource problems if the number of available resources for Air/Ground/Air assignments is cut by 14% (i.e. 1000 of 7000 channels).

These trials were conducted in the scope of defining algorithms for NUFAS 3.

Based on the trial results, the FMB Staff is developing a new weighting factor algorithm to take into account the availability of frequency spectrum for assignments. The new algorithm is still to be tested.

11 Conclusions

It is concluded that :

a. Ordering assignment requests is essential in a bulk assignment process.
b. The attributes of the FMSC frequency records are so complicated that the optimal weighting algorithm can only be determined empirically.
c. An improved result might be achieved by taking into account frequency spectrum availability regarding requests

12 Way Ahead

The future NUFAS 3 system will use an improved weighting factor algorithm. Further trials of the impact of the adjacent channel protection criteria are also planned.

Figure 1: Assignment Number as function of Frequency

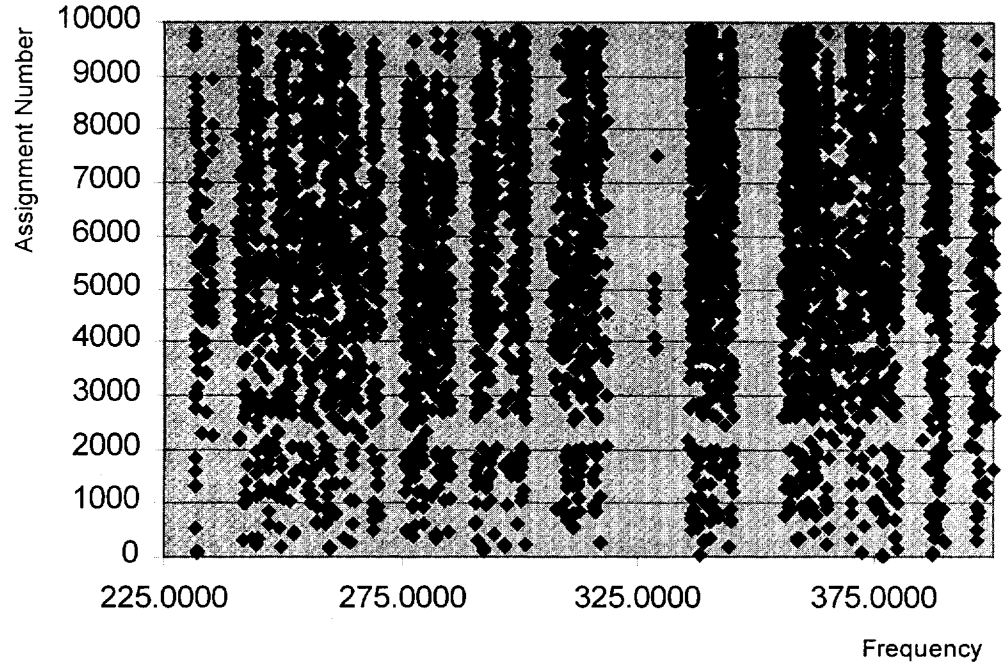


Figure 2: Resource Reuse Rate as function of the Resource Number

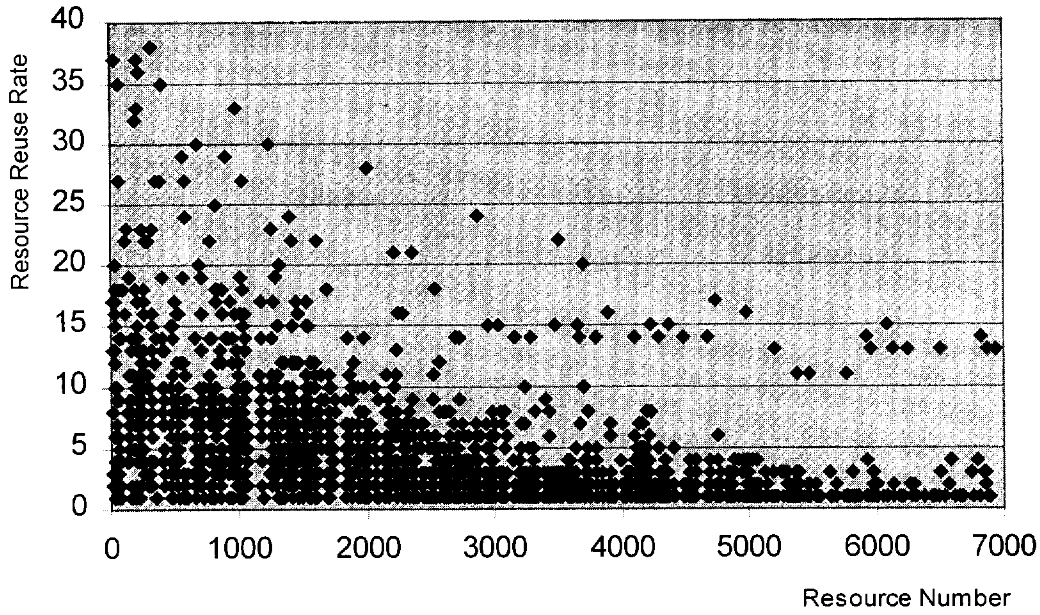
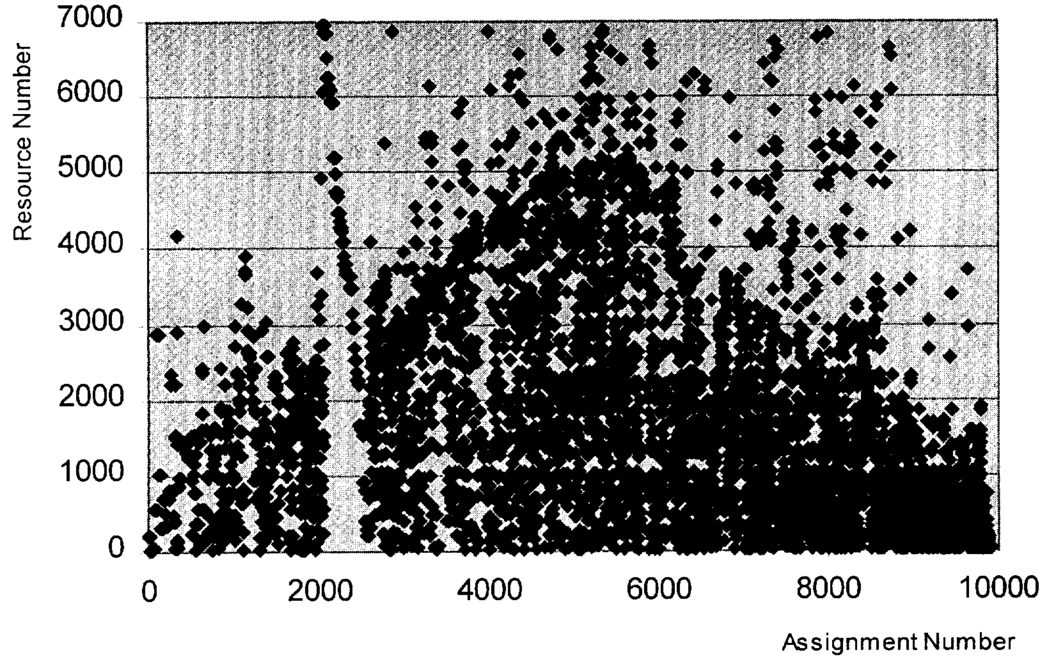


Figure3: Resource Number as function of Assignment Number



PAPER TITLE : **Paper 10 - The Impact of Protection Criteria Assignment Order on Military Air Ground Air Frequencies**

AUTHOR : K. S. Kho, M Elliott

NAME : N. Wheadon

QUESTION :

1. What form are the requests received at the FMB, using standard templates; manually or automatically using e-mail?
2. How do you account for Frequency Hopping (FH) system and the possible re-use of those frequencies?

ANSWER :

1. Unfortunately, for the moment our message exchange is conducted manually using the so-called ABNF Forms or 14 point message. However, in several months time, it will be automatic. The messages will be electronic messages under the so-called ARFA data Exchange Format (ADEF) supported by a user interface called ARCADE.
2. The management of Frequency Hoppers in the UHF band is treated separately from the Fixed Frequency radios. The UHF frequency hoppers Have Quick II and SATURN are not managed using NUFAS but a program called NATO Net Number Assignment System (NANNAS). NANNAS assigns net numbers to frequency hoppers. We have only a single hopset for UHF frequency hoppers. Therefore, we do not assign hopsets (only net numbers).

Additional remarks by Co-author :

3. The processing time for a complete bulk assignment is several hours, and so we have assumed that an optimising algorithm may be impractical.
4. Also I would like to clarify that the wartime hopsets for the current UHF frequency hopping radio, HAVE QUICK II, are quasi-fixed in the radio. This is because otherwise it is difficult to ensure interoperability between aircraft coming together from different countries for a common mission : it is necessary for interoperability to have exactly the same frequencies stored in the radio. Although the hopsets are held in re-programmable memory there is no infrastructure available to change to a new hopset. However, for the next generation radio SATURN there will be an infrastructure to make it practical to change hopset, although this will still require a significant lead time. These hopsets should therefore be considered as allotments from the NATO Frequency Management Sub-Committee, subject to only occasional change. The hopping channels however, are analogous to frequencies and are managed by the net-number assignment tool already mentioned.

PAPER TITLE : **Paper 10 - The Impact of Protection Criteria Assignment Order on Military Air Ground Air Frequencies**

AUTHOR : K. S. Kho, M. Elliott

NAME : E. Tsang

QUESTION :

1. Could you explain where the algorithm spends the processing time? (This is to follow up Glyn Wyman's question.)

ANSWER :

1. The bulk assignment process is divided in 5 phases. The phases are as follows:

   Phase 1 : Validation of the records
   Phase 2 : Processing of Pre-assigned records
   Phase 3 : Construction of Matrix
   Phase 4 : Weighing
   Phase 5 : Assignment of frequencies

2. The most time consuming process is the construction of the matrix. The other processes are also time consuming but require much less than the time required for the matrix construction phase.

NAME : G. Wyman

QUESTION :

1. Have you considered optimising the sequential algorithm applying multiple attempts rather than improve the initial order with a single pass?

ANSWER :

1. Yes, we have considered optimising the algorithm. In the next system we will try to improve the results with other criteria in selection of valid frequencies e.g the best D/U ratio, least re-use etc. Our problem is more in obtaining accurate data from the users. The time sharing aspect for instance is still not yet exploited. The bulk assignment process is a very long process. With our current machines, the process last 48 hours. In the future, with the new machines, it is expected that it would be possible to apply further optimation algorithms.

PAPER TITLE : **Paper 10 - The Impact of Protection Criteria Assignment Order on Military Air Ground Air Frequencies**

AUTHOR : K. S. Kho, M. Elliott

NAME : Mr. Jaeger

QUESTION :

1. The UHF frequency band seems to be very occupied. In the previous presentation, we heard that a new UHF radar is in development. Have the possible effects of the introduction of this radar already been considered?

ANSWER :

No, this radar is still not yet considered. In NATO, we have a procedure called "Frequency Supportability Request" which is similar to the US procedure we heard about in an earlier briefing. All electronic equipment deployed in NATO Europe must have frequency supportability. The introduction of radar in the UHF band will cause problems. It is still not yet possible to say more about this since we have not received any frequency supportability request for UHF radars. However, if we do receive a frequency supportability request for UHF radars, we will then consider it carefully.

ARE THERE EFFECTIVE BINARY FREQUENCY SEPARATION CONSTRAINTS FOR FREQUENCY ASSIGNMENT COVERAGE PROBLEMS?

Joe Bater, Peter Jeavons, David Cohen and Nick Dunkin
Department of Computer Science
RHUL, Egham, Surrey, TW20 0EX, England
email: {joe,pete,dave,nickd}@dcs.rhbnc.ac.uk

SUMMARY

Frequency assignment in cellular radio networks is often modelled using *binary* constraints between pairs of transmitters. These constraints restrict possible frequency channel assignments in order to reduce the risk of unacceptable interference. Here we question the adequacy of binary constraints of this kind to represent the problem effectively.

In this paper we examine the use of binary constraints, based on predicted interference between pairs of cells [1, 2, 3]. Having generated the binary constraints, they are solved using standard heuristic solution techniques (e.g. [4]), and tested back against the original system model.

In many of the instances these solutions fail to provide complete coverage. By considering solutions calculated directly from the system model (i.e. a single global constraint), we obtain solutions to all problem instances which provide perfect predicted coverage.

To achieve coverage in the binary model it is necessary to tighten the constraints in these instances by increasing the reference carrier-to-interference thresholds at which constraints are enforced. This additional restriction forces the number of channels used in the solution to grow beyond that required of the global constraint solution.

We conclude that representing the frequency assignment problem using binary constraints may be inadequate to capture the essential features of the problem, and that it may be necessary to include *higher-order* information in any effective model.

Another conclusion concerns the lower bounds on the number of channels required that are sometimes calculated from a binary constraint model, using maximal cliques in the constraint graph, or other techniques [5, 6]. If the binary constraint model is not an adequate representation of the problem, then any such lower bound may be unreliable as an estimate of the spectral resources required for a cellular radio system.

1. INTRODUCTION

The frequency assignment problem (FAP) is a genuine hard problem of practical importance to both civil and military administrations. The FAP can, according to it's particular definition, encompass one of a number of possible tasks. We are particularly concerned with coverage problems in dense networks, such as those existing for mobile telephone networks.

The essential characteristic of these is that a frequency must be assigned to each transmitter (Tx) in such a way that the carrier-to-interference ratio (C-I) at any receiver (Rx), receiving signal from a specified Tx, in the region to be covered is greater than some specified acceptable value.

$$\forall \mathrm{Rx}, \exists \mathrm{Tx}_{\mathrm{link}} : \text{C-I} \geq \text{C-I}_{\mathrm{acceptable}} \quad (1)$$

Coverage tasks can take a number of guises. Two common such coverage-type tasks are:

1. Minimise the spectrum requirement, by finding a satisfactory assignment which uses the smallest number of distinct frequencies.
2. Maximise the solution quality, by finding an assignment using a given fixed allocation of frequencies that provides the best possible coverage (there are many possible optimisation criteria).

The work described here concerns the first of these tasks and investigates some of the assumptions commonly made when modelling this kind of FAP. Specifically, we question the assumption made in many previous studies that binary constraints (constraints which restrict the values on pairs of transmitters) can be efficiently found which specify the problem effectively [1, 7, 8, 9, 10, 11, 12, 13].

To carry out the study described here, a small network of transmitters and receivers was modelled, emulating the system characteristics of a mobile telephone network (GSM). Binary constraint solutions were generated using a standard technique and compared to the solutions ob-

Paper presented at the RTO IST Symposium on "Frequency Assignment, Sharing and Conservation in Systems (Aerospace)", held in Aalborg, Denmark, 5-7 October 1998, and published in RTO MP-13.

tained by expressing the original system requirements as a global constraint.

2. THE MODEL EMPLOYED

In the land-mobile service TDMA radio networks are typically designed with a cellular structure in which each transmitter provides the signal for all receivers in its neighbourhood [14, 15, 16].

In previous work ([17]), we generated a model containing a fixed set of Tx sites placed at random (uniform distribution) on a 100 by 100 matrix of Rx sites which is referred to as "Pinkville" (such an approach has been used before e.g. [18, 19]), in the experiments described here we use a similar modelling process on a number of instances ("Pinkville1", "Pinkville2" and "Pinkville3" - collectively referred to as the "Pinkville" instances). Each Rx position was "scaled" by a factor of 50, Pinkville1 is thus notionally a 5km by 5km region. We randomly generated 50 Tx locations, and then generated problem instances with n transmitters by simply using the first n of these Tx locations (for n in the range 20-50). The positions of the transmitters in Pinkville1 are shown in Figure 1.

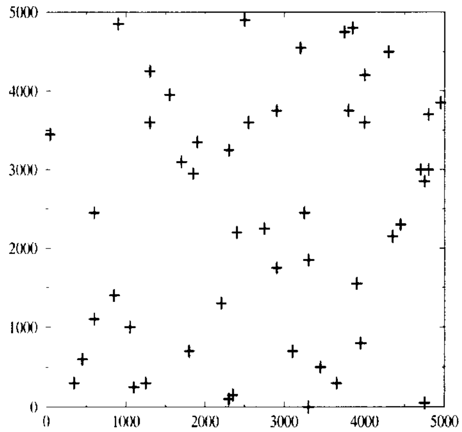


Figure 1: Positions of transmitters in the 50Tx Pinkville1 instance after "scaling"

The system model used was as follows:

1. Tx were assumed to operate with uniform effective radiated power. That is, all transmitting powers were equal, and propagation was omnidirectional

2. Open-space loss was assumed to be the only cause of signal attenuation. This was modelled using an inverse fourth power law:

$$P_{\mathrm{Rx}} = \frac{P_{\mathrm{Tx}}}{d^4}. \quad (2)$$

3. Cells were defined by assuming that the desired signal at each Rx was received from the nearest Tx (equidistant Tx were chosen arbitrarily).

4. Frequencies were assumed to be chosen from a contiguous block of discrete channels. Adjacent channel rejections followed the GSM specification [20, 21]:

offset	interference	rejection
0	C/Ic	(none)
1	C/Ia_1	-9dB
2	C/Ia_2	-41dB
3	C/Ia_3	-48dB
≥ 4		(total)

5. Multiple interfering signals at the same receiver were combined using power-summation:

$$\text{Total Interference} = \textstyle\sum_{i=1}^{n} \text{Interference}_i$$

3. CHOOSING BINARY CONSTRAINTS

Although the various specifications of radio services suggest no particular approach to choosing frequencies [20, 21, 10, 2], a graph theoretic approach is commonly used. When the problem is tackled in this way, the requirements on the desired assignment are described by constraining the frequencies that may be assigned to certain pairs of transmitters. These constraints typically take the form:

$$|f(\mathrm{T}x_i) - f(\mathrm{T}x_j)| \geq k \quad (3)$$

meaning that the frequencies assigned to Tx_i and Tx_j are required to be $\geq k$ channels apart in order to avoid excessive interference at any Rx.

In this approach, the FAP is considered to be analogous to GRAPH K-COLOURING, GRAPH T-COLOURING or GRAPH LIST-COLOURING [8, 9, 11, 12, 18, 19, 22, 23]. Models more involved with realistic data have raised concerns on this issue [7, 24], though their response is often to reformulate the problem as an optimisation problem involving *hard* and *soft* constraints (the latter of which express the optimisation criteria on a now *partial CSP*).

The fact that different approaches have been explored raises the question of how the original requirements for an adequate C-I at each Rx may be translated into these binary constraints between pairs of transmitters. This problem is particularly acute when attempting to find solutions which minimise the number of frequency channels used, because the choice of binary constraints must then satisfy two opposing criteria:

1. MINIMUM SPECTRUM The binary constraints chosen must be sufficiently loose. If they are too restrictive, then the solutions found will use an unnecessarily large number of frequency channels.
2. ADEQUATE COVERAGE The binary constraints chosen must be sufficiently tight. If they are not restrictive enough, then at least some of the solutions found will not satisfy the coverage requirements.

The binary constraint representation of FAP CPs commonly takes one of two forms:

1. *frequency-distance (f*d) constraints* - (e.g. [10, 12, 18, 25, 26], constraints are based on "re-use" distances, which specify, under a simplified geographic model), the minimum distances Tx must be apart so as not to cause excessive interference on a co-channel assignment.
2. *frequency-separation (f*s) constraints* - (e.g. [1, 2, 3]), which are based on interference predictions. Consider the interference on the coverage areas of two cells both tentatively assigned co-channel. These constraints then specify the required frequency separation to avoid interference (these constraints might thus be more accurately described as "interference avoidance constraints").

In our experience, the data sets made available to many researchers are often already in binary constraint form, so this question has not received the attention it deserves.

In a previous paper we established that the simple *f*d* method does not lead to satisfactory results on the type of networks we are considering [17]. Therefore, in this report we specifically concentrate on *f*s* type constraints.

We generate these by comparing the interference on a pair of cells by their opposing Tx (See Figure 2).

Assuming that both are given a co-channel assignment the C-I at each Rx in each cell is checked. If this proved to be adequate then no *f*s* would be necessary (both Cell A and B have regions marked 'n' for no constraint), alternatively the interference level could require that assignments be separated by up to ±3 channels (Cell A has regions which require that Tx_B is at least

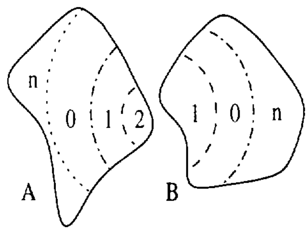


Figure 2: Constructing *f*s* constraints

2 channels separate, whereas Cell B only needs to be at least 1 channel apart from Tx_A.

Given that binary constraints are inherently symmetric, the most restrictive constraint is always chosen if there is asymmetry. Thus, in the example of Figure 2 a frequency-separation constraint of at least 2 channels would be applied.

4. RESULTS

4.1. Experiment One

For each of the Pinkville1 problems, instances of 20, 30, 40 and 50 Tx were generated. We then generated simple *f*s* constraints and solved these using the software package *FASoft* [27] (FASoft is a suite of binary constraint meta-heuristic solution techniques for the discrete channel FAP. It was developed by the University of Cardiff and Pinkville1 in association with the UK Radiocommunications Agency [27, 28].) Here FASoft was used with a sequential algorithm, whose result is then improved using Tabu Search (TS). (It should be noted that these binary constraint solutions were always at or near the predicted lower bounds on the span derived from the binary model - also calculated by FASoft [27, 28]). A solution at these lower bounds was very quickly found, but when the coverage was checked against the original model it was found not to give complete coverage (See Figure 3 and Table 1.

It would be a fair assessment that the predicted coverage of the simple binary constraint solutions was not bad, but for the purposes of this investigation we want to see how effective the *f*s* model could be at obtaining solutions with 100% predicted coverage. Since the initial settings are too loose (solutions to the binary constraints do not translate to coverage problem solutions) we now make the binary constraints more restrictive in order to achieve full coverage solutions.

4.2. Experiment Two

In the second experiment we attempted to generate *f*s* constraints to guarantee 100% 9dB cov-

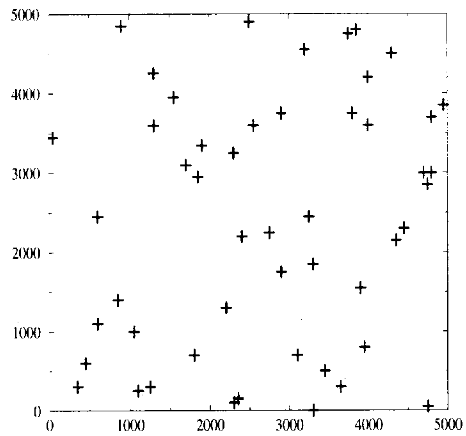


Figure 3: Coverage (to 9dB C-I) of the simple f^*s constraint solution. The dots represent Rx receiving suitable quality signal, the absence of such a dot indicates a "hole" in coverage

Binary				
#Tx	20	30	40	50
#f	6	6	6	5
9dB cov	88.1%	86.9%	88.1%	86.1%
12dB cov	71.5%	72.6%	74.5%	70.5%
15dB cov	55%	58.3%	58.2%	55.6%
Global				
#Tx	20	30	40	50
#f	9	12	14	14
12dB cov	98%	97.7%	97.9%	98.6%
15dB cov	92%	90.9%	91.7%	93.7%

Table 1: Coverage of simple f^*s and global constraints on Pinkville1 instances

erage. Using the same method to assess interference between pairs of cells, we altered the threshold C-I sensitivity at which a constraint should be enforced. This was done by universally adding a PRUDENCE value (in dB) to the GSM specified reference sensitivity levels.

In Figure 4 the calibration process by which additional C-I sensitivities were considered for constraint enforcement are shown. To guarantee full coverage on Pinkville1 it was necessary to set constraints for an additional 17dB (that is it was necessary to generate f^*s constraints for a notional C-I value of 26dB to achieve coverage at 9dB).

Clearly this additional restrictiveness will have an effect on the required frequency span. The growth of span according to the PRUDENCE factor used is shown in Table 2. In the course of these

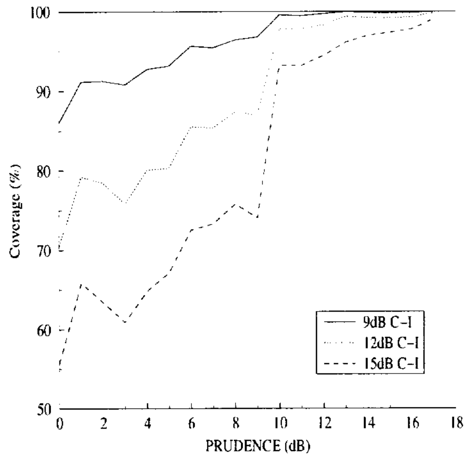


Figure 4: Coverage of the 50 Tx Pinkville1 instance

experiments a considerable jump in coverage occurred after a PRUDENCE factor of 10dB was employed (See also Figures 4 and 5). Given this extra restrictiveness in constraint setting coverage at the 12dB and 15dB level were observed to rise dramatically across all three 50Tx Pinkville instances.

In Table 2 *global* constraint solution spans are compared to those of the *modified binary* constraints for the Pinkville1 instance (the required PRUDENCE factor to guarantee 100% coverage is also shown).

#Tx	Global	Binary	PRUDENCE
20	9	11	14dB
30	12	15	15dB
40	14	18	17dB
50	14	21	17dB

Table 2: Spans of 100% coverage solutions. These are expressed as global constraint solutions (upper bound) and "prudent" binary f^*s constraint solutions (lower bound) on Pinkville1 instances. The required PRUDENCE restrictiveness factor is also shown (expressed in dB)

5. CONCLUSIONS

The f^*s constraints present a more interesting and subtle picture, and appear to be much more successful than the less sophisticated f^*d constraints which had been considered in previous work [17]. However, the simple binary f^*s constraint solutions appear to inherently produce

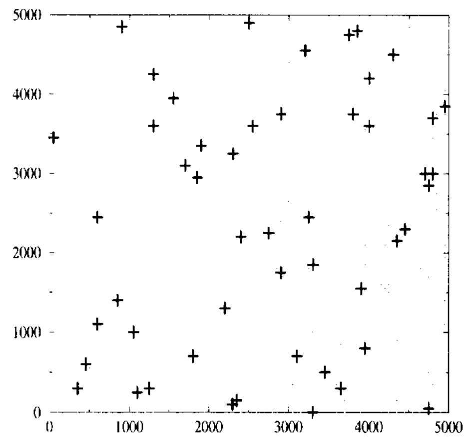


Figure 5: Coverage of the 50 Tx Pinkville1 instance with 10dB PRUDENCE applied

black-spots in actual coverage due to the effect of multiple interferences. The simple f^*s formulation cannot adequately account for multiple interference from sets of local Tx.

It is widely believed that "sensible" settings for binary constraints should provide an adequate model for frequency planning. In the second experiment we showed that it is hard to entirely eliminate coverage black spots when using binary constraints, without making these increasingly restrictive and hence upping the span drastically. In short, to guarantee coverage these "sensible" settings had to be quite "extreme". Thus would we could argue for the use of global constraints on the grounds of achieving better quality solutions (even if only at first in the solving of particularly troublesome regions).

One aspect of this conclusion concerns the lower bounds on the number of channels required that are sometimes calculated from a binary constraint model, using maximal cliques in the constraint graph, or other techniques [5, 6]. If the binary constraint model proves an inadequate representation of the problem, then any such lower bound may be unreliable as an estimate of the number of frequency channels necessary to provide adequate coverage. Using such lower bounds may therefore result in an unreliable estimation of the spectral resources truly required for cellular radio.

The authors are currently investigating the possibilities of *higher-order* constraint models for frequency planning. Such models can be produced using constraints involving more than two transmitters at a time. Early results on a regular grid system point to increased flexibility and improved coverage without the expense of largely increased span.

References

[1] A. Knälmann and A. Quellmalz. Solving the frequency assignment problem with simulated annealing. In *Proceedings of the Ninth International Conference on Electromagnetic Compatibility*, pages 233–240, 1994.

[2] R.G. Strużak. Optimum frequency planning: a new concept. *Telecommunications Journal*, 49:29–36, 1982.

[3] D.A. Youngs. Frequency assignment for cellular radio networks. In C. Mitchell, editor, *Applications of Combinatorial Mathematics*, number 60 in IMA Conference Series, pages 233–241, 1995.

[4] C.R. Reeves. *Modern Heuristic Techniques for Combinatorial Problems.* John Wiley & Sons, New York, NY., 1995.

[5] N.W. Dunkin and S.M. Allen. Frequency assignment problems: Representations and solutions. Technical Report CSD-TR-97-14, Royal Holloway, University of London & University of Glamorgan, June 1997.

[6] Hurkens, C.A.J. and Tiourine, S.R. *Upper and Lower Bounding Techniques for Frequency Assignment Problems*, 1995. EUT Internal Report for EUCLID Project RTP 6.4.

[7] J.F. Boyce, H.D. Dimitropoulos, G. vom Scheidt, and J.G. Taylor. GENET and Tabu search combinatorial optimisation problems. In *World Congress on Neural Networks (WCNN95)*, Washington DC, 17-21 July 1995.

[8] G.R. Bradbeer. Graph colouring for frequency assignment of combat radio. Defence Evaluation and Research Agency Unpublished Technical Report, December 1996.

[9] W. Crompton, S. Hurley, and N.M. Stephens. Applying genetic algorithms to frequency assignment problems. *Neural and Stochastic Methods in Image and Signal Processing III*, 2304(28):76–84, 1994.

[10] W.K. Hale. Frequency assignment: Theory and applications. *Proceedings of the IEEE*, 68(12):1487–1514, December 1980.

[11] A. Kapsalis, V. Rayward-Smith, and G.D. Smith. Using genetic algorithms to solve the radio link frequency assignment problem. In *Proceedings of the 2nd international conference on artificial neural networks and genetic algorithms*, pages 37–40, 1995.

[12] R. Leese. Tiling methods for channel assignment in radio communication networks. In *Proceedings of the 3rd International Conference on Industrial and Applied Mathematics*, 1995.

[13] D.H. Smith, S. Hurley, and S.U. Thiel. Improving heuristics for the frequency assignment problem. *European Journal of Operations Research*, To Appear.

[14] M.S. Appleby. The UK cellular system. In Macario, R.C.V., editor, *Personal & Mobile Radio Systems*, chapter 9, pages 192–207. Peter Peregrinus, Stevenage, United Kingdom, 1991.

[15] D.M. Balston. The pan-European cellular technology. In Macario, R.C.V., editor, *Personal & Mobile Radio Systems*, chapter 14, pages 290–319. Peter Peregrinus, Stevenage, United Kingdom, 1991.

[16] G.D. Lochtie and M.J. Mehler. Subspace approach to channel assignment in mobile communications networks. *IEE Proceedings on Communications*, 142(3):179–185, 1995.

[17] J.E. Bater, P.G. Jeavons, and D.A. Cohen. Non-binary modelling of frequency assignment leads to significant reduction in predicted spectrum requirements. In J.M. Janiszewski, W. Moroni, and W. Sega, editors, *14th International Wroclaw Symposium on Electromagnetic Compatability (EMC'98)*, pages 705–709, June 1998.

[18] L.A. Berry. Potential contribution of optimum frequency assignment to efficient use of the spectrum. In *Proceedings of the IEEE International Symposium on Electromagnetic Compatibility*, pages 409–412, 1990.

[19] J.A. Zoellner and C.L. Beall. A breakthrough in spectrum conserving frequency assignment technology. *IEEE Transactions on Electromagnetic Compatability*, EMC-19:313–319, 1977.

[20] ETSI. *European digital cellular telecommunications system (Phase 2); Radio network planning aspects (GSM 03.30)*, February 1995.

[21] ETSI. *Digital cellular telecommunications system (Phase 2); Radio transmission and reception (GSM 05.05)*, November 1996.

[22] E. Malesińska. List colouring and optimization criteria for a channel assignment problem. Technical Report 458/1995, Technische Universität Berlin, 1995.

[23] J. Žerovnik. Experiments with a randomized algorithm for a frequency assignment problem. Technical Report 97-27, Ecole Normale Supérieure de Lyon, September 1997.

[24] G. Ottosson and M. Carlsson. Using global constraints for frequency allocation. Technical Report TR-97-07, ASTEC, 1997.

[25] M.C. Delfour and A. De Couvreur. Interference-free assignment grids - part i: Basic theory. *IEEE Transactions on Electromagnetic Compatability*, 31(3):280–292, August 1989.

[26] D.H. Smith, S. Hurley, S. Thiel, and S.M. Allen. Benchmarks for the channel assignment problem. *IEEE Transactions on Vehicular Technology*, Submitted.

[27] S. Hurley, D.H. Smith, and S.U. Thiel. FASoft: A system for discrete channel frequency assignment. *Radio Science*, 32(5):1921–1939, 1997.

[28] D.H. Smith, S. Hurley, and S.U. Thiel. *FASoft 2.0: User Manual*. University of Glamorgan and Cardiff, April 1997.

PAPER TITLE : **Paper 11 - Are There Effective Binary Frequency-Separation Constraints for Frequency Assignment Coverage Problems?**

AUTHOR : J. Bater et al

NAME : E. Tsang

QUESTION :

You have demonstrated that a model using binary constraints did not allow good solutions (compared to a model using global constraints). But you can't rule out the possibility of finding good binary models, can you?

ANSWER :

No I can't. In fact it is obvious that an optimal set of binary constraints exists. Simply take a known optimal solution and project the assignments onto pairs of variables and set these as constraints. This method has obvious problems.

It is exactly the *a-priori* setting of binary constraint relations that we feel may be more difficult than previously thought. The constraint model must take into account two conflicting requirements (page 1), in this scheme a sliding scale of "good" and "bad" constraints can be easily visualized. The question is then: is there a good setting somewhere in the middle, which satisfies both criteria? From the work described here, which increases the stringency of setting constraint relations, we argue that there might well not be.

How else might constraints be set? One answer is to use higher-order constraints (HOCs). These constraints find suitable relations over sets of variables (transmitters) larger than two. Methods utilizing higher-order constraints are well known in the constraint community and have led to advances in other fields, such as scheduling.

A HOC applied to a single central cell would constrain the permitted assignments to Tx in the neighbouring region which are deemed to have potential interference effects. For example, if cell 5 was anticipated to have risk of interference from Tx 1,12,17,21,24,34 and 43 all of these Tx could be grouped into the scope of a HOC (in the normal binary model there would be an edge between Tx5 and these other Tx). A relation of permitted assignments generated with multiple interferences directly taken into account naturally and concisely captures possible assignments to this scope.

Scope	5	1	12	17	21	24	34	43
Relation	0	1	2	2	2	2	1	2
	0	2	2	1	2	2	1	1
	0	2	2	2	2	1	1	2

Note that the HOC can have a number of labellings, each of which describe possible simultaneous settings of assignments around cell 5 which guarantee adequate C-I on the Rx of that cell.

We have successfully implemented higher-order constraint on the Pinkville instances and find that their direct ability to account for multiple interference and increased flexibility lead to improved results on the simple binary F*S and PRUDENT binary F*S constraint models.

These HOCs can be broken down into smaller constraints, possibly binary ones. Observe that all three labelling on Tx 12, 21 and 34 each have the same restriction. These can be removed from the HOC and replaced using binary constraints only. The remainder of the HOC contains restrictions that allow trade-offs between interference contributions might be decomposed using techniques available from the constraint satisfaction research community.

EMERGING TECHNIQUES FOR DYNAMIC FREQUENCY ASSIGNMENT: MERGING GENETIC ALGORITHMS AND NEURAL NETWORKS

C. Bousoño-Calzón
A.R. Figueiras-Vidal
Dpto. Tecnologías de las Comunicaciones
Univ. Carlos III de Madrid
C/Butarque, 15; 28911 Leganés (Madrid), Spain
Tel: +34.91.624.9952, Fax: +34.91.624.9430, E-m: cbousono@ing.uc3m.es

SUMMARY

Genetic algorithms and neural networks have previously been applied to the hard problem of assigning channels in mobile communication systems. The interest in these algorithms is due to their generality and the possibility of fast hardware implementations that adapt to dynamical environments. Nevertheless, these algorithms perform differently: neural nets are known to better satisfy allocation constraints while genetic algorithms allow for global optimisation. We propose here to merge the best features of both algorithms in a quite natural manner. Simulations show that this merging has good performance, and suggest a new interesting direction for research.

1 INTRODUCTION

Along with increasing demand, new mobile services require more bandwidth, and suffer large traffic fluctuations during sessions. The dynamic allocation of channels can in part alleviate this tendency, but the kind of problems involved in such management are known to be quite hard [Hale], especially if real-time solutions are required to follow changing environment conditions.

Emerging techniques, such as Genetic Algorithms (GAs) [Lain] or Neural Networks [Kunz, Funabiki] among others, provide an interesting approach to these problems. First, they can be easily applied to quite different combinatorial optimisation problems without developing ad hoc algorithms (such as [Kim]). These algorithms also allow cheap and fast hardware implementations [Haykin].

In order to make use of these algorithms, a representation technique for the possible solutions of the problem is required; that is, we need to formulate a search space for the algorithm. It is important to note that this formulation is strongly related to the constraints of the problem at hand. The usual approach taken to tackle these constraints is the so-called "Penalty Method". It consists of using a search space larger than the space of valid solutions, and inserts the constraints of the problem as additive terms in the cost function.

However, this approach has some drawbacks. First, it is difficult to ensure that the solution obtained is a feasible assignment, i.e., that it be able to fulfil the constraints of the allocation. Second, even if valid solutions are obtained, the constraint terms added to the cost function makes it difficult for the algorithm to find optimal solutions.

Since it is well known that some of the emerging algorithms are better at satisfying the constraints, and others are better at finding better cost solutions, we propose a hybrid algorithm which combines two emerging techniques. Specifically, we use a neural network to satisfy the constraints along with a genetic algorithm to find global optimals. The design of the neural network has the main objectives of providing feasible solutions and improving the convergence properties of the network. In addition, the synthesis of the genetic algorithm is simplified with this approach, since the information provided by the cost function to the algorithm is not perturbed by the constraint terms.

We first introduce a classical mathematical formulation for the general problem of assigning channels (in this paper, we will use "channel" and "frequency" interchangeably). In Section 3, the design of the neural network is given. The hybrid algorithm is proposed after briefly revising a basic form of GA in Section 4. Section 5 presents the simulations and a comparative evaluation of the algorithm. Finally, Section 6 provides some conclusions and proposes some further research on this topic.

2 PROBLEM FORMULATION

Among the multiplicity of problems related to the assignment of channels in a mobile communication system, we have selected a general form of allocation by Gamst and Rave for arbitrary inhomogeneous networks [Funabiki]. The application of the algorithm given in this paper to any other form of frequency allocation is quite easy. We illustrate here how to include constraints by considering two problems, one of them being a generalisation of the other, and finding some interesting trade-offs from their comparison.

The m channels of the system will be represented by a number in $\{0,1, ..., m\}$. The electromagnetic compatibility constraints in a n-cell mobile network are given by a n x n symmetric matrix, which is called the compatibility matrix, C. Each row and column stands for a cell, and element c_{ij} is an integer number which indicates that if the channel p is assigned to cell i and the channel q is assigned to cell j, then $|p-q| \geq c_{ij}$. Thus, $c_{ij} = 1$ means a co-channel constraint between cells i and j. An adjacent channel constraint between the same cells can be represented by $c_{ij} \geq 2$, and the co-site constraints are represented by the diagonal terms.

We can also take into account a minimum number of frequencies requests in a cell through a vector R, where r_i represents the number of channels solicited to the system by cell i.

The first problem we consider, P1, involves maximising the number of channels in the system given the compatibility matrix C. The second problem, P2, is similar to P1, except that we add by-cell demand constraints (through vector R).

2.1 The 0-1 formulation.

It is traditional to represent the allocation of channels to cells using an m x n matrix F of 0's and 1's, where $f_{ci}=1$

Paper presented at the RTO IST Symposium on "Frequency Assignment, Sharing and Conservation in Systems (Aerospace)", held in Aalborg, Denmark, 5-7 October 1998, and published in RTO MP-13.

means that channel c has been assigned to cell i, and f_{ci}=0 indicates no assignment between these cells.

With this representation, we can mathematically formulate the problems as:

(P1)

$$\max \sum_{c,i} f_{ci}$$

such that

$$|f_{pi} - f_{qj}| \geq c_{ij}, \quad \forall p,q,i,j \;\; s.t. \;\; f_{pi} = f_{qj} = 1$$

$$f_{pi} \in \{0,1\}, \quad \forall p,i$$

(P2)

$$\max \sum_{c,i} f_{ci}$$

such that

$$|f_{pi} - f_{qj}| \geq c_{ij}, \quad \forall p,q,i,j \;\; s.t. \;\; f_{pi} = f_{qj} = 1$$

$$\sum_{p} f_{pi} \geq r_i, \forall i$$

$$f_{pi} \in \{0,1\}, \quad \forall p,i$$

3 THE NEURAL NETWORK

The neural network presented below is a Hopfield type net similar to that in [Funabiki]. However, it has some new properties which makes it more appropriate to the hybrid algorithm of Section 4. It has been designed to always converge, and the number of iterations to reach convergence has also been decreased by a factor of three. A more detailed comparison between our network and Funabiki's is provided in Section 4.

We first introduce the general Hopfield model, and then we give the equations for solving the problems of Section 2.

3.1 The Hopfield model for (P1).

The original model was introduced by Hopfield in 1982 [Hopfield]. It consists of a set of variables $\{V_i\}$, called neurons, usually arranged in a matrix structure and interconnected in such a way that, once initialised to a certain value, the system relaxes to a point which can be considered to be a solution to the problem. Since the introduction of this model, there have appeared numerous variations to this basic scheme, but the essential problem still remains one of choosing the structure so that the solution obtained is a feasible one[1].

[1] For a good introduction to this subject see [Hertz] and [Cichocki].

The behaviour and structure of the neural net can be described by the following set of equations:

$$\partial_t V_{ij} = -\partial_{V_{ij}} E \cdot V_{ij}(1 - V_{ij}), \;\; \forall i = 1...m, j = 1...n$$

where V_{ij} corresponds to f_{ij} in the problem representation (see Section 2.1), the V_{ij} and $(1\text{-}V_{ij})$ terms constrain the values of the neurons to the interval (0,1), and E is the so called "Energy function" which stands for the constraints given by the compatibility matrix:

$$E_1 = \sum_{p=1}^{m}\sum_{i=1}^{n} \sum_{\substack{q=max(1,p-c_{ij}+1) \\ q \neq p}}^{q=min(m,p+c_{ij}-1)} \sum_{\substack{j=1 \\ c_{ij}>0}}^{n} V_{pi}V_{qj}$$

plus the cost function:

$$E_2 = \sum_{p=1}^{m}\sum_{i=1}^{n} V_{pi}$$

That is, $E=E_1+\gamma E_2$, where γ is a weight factor. The dynamics of such a system is restricted to the inner part of a hypercube, $(0,1)^{nm}$, with one dimension per neuron. The evolution of the network can be described by the minimisation of the Energy function (which is a Lyapunov function for the system, see [Hertz] for details) which goes to E=0.This means that a feasible solution to the problem has been found. Sometimes the network gets stopped at a point of the inner part of the unit hypercube, such a problem can be easily circumvented by adding some noise to the gradient of the net (since the stagnation points are known to be unstable).

It should be noted that every initialisation of the network is (ideally) associated with a solution of the problem; however, the noise added to avoid the stagnation points can eventually map different solutions to the same initial point.

3.2 The cell-by-cell demand constraints (P2)

The standard approach for including the request constraints for P2 is to add an extra term to the Energy function such as [Funabiki]:

$$E_3 = \sum_{i=1}^{n}(\sum_{p=1}^{m} V_{pi} - r_i)^2$$

This term contributes to the gradient of the neural net as follows[2]

$$\partial_{V_{pi}} E_3 = \sum_{q=1}^{m} V_{qi} - r_i$$

However, this procedure assumes that it is possible to serve all the requests. The fact that this cannot be guaranteed in a dynamical mobile system means that the neural network may not converge. Even if the demand

[2] We are not taking into account the weight factor here.

In order to avoid the previously mentioned problems, we propose a new term:

$$\partial_{V_{pi}} E_3 = \mathrm{sgn}\left[\sum_{\substack{q=1 \\ q \neq p}}^{m} \mathrm{sgn}(V_{qi} - 1 + \varepsilon) - r_i\right] - 0.5$$

where the "sgn" function has value 0 if the argument is less than 0, and 1otherwise, and ε is a small term added to ensure that the evolution of the network is strictly inside the unit hypercube.

The meaning of this term is that only the assigned frequencies in cell i (i.e. those neurons already converged to 1) are counted. If its number is greater than the requests made in cell i, then neuron Vpi tends to decrease since the demand is already satisfied. However, if this number is less than the number of requests, the neuron tends to increase.

Although this term is better than the previous for obtaining feasibility, the quality of the obtained solution is worse. We can do better if we use both heuristics, that is, if the net is running with Funabiki's heuristic up to a point (we have made it equal to iteration 25th), and then it is switched to the proposed heuristic until it converges to a feasible solution (see Section 5 for comparisons and results). Some other trade-offs can be made.

4 THE HYBRID ALGORITHM

The previous sections have tackled the assignment problem from the neural network perspective. Although the net is able to retrieve feasible solutions, these are just locally optimal. The Genetic Algorithms (GA) have also been applied to the assignment problem (see for instance [Lai]), but although they can search for global optimals, they do not yield feasible solutions.

In this section, we combine both methodologies in a natural manner. Comparative results are given in the next section.

4.1 The genetic algorithm

We take the basic scheme for these algorithms given by Goldberg in [Goldberg]. It consists of a population of 50 genes (a gene is a string of 0's and 1's of length k) which evolves through the application of three simple genetic operators: reproduction, crossover and mutation.

The reproduction follows the roulette wheel rule. Each gene is given a real number which is the fitness for the problem, that is, the number of the assigned channels. In a second step, it is copied to a new population with a probability that is given proportional to its fitness (normalised by the sum of all the fitness values of the population).

To apply the crossover operator, each gene in the reproduced population is mated at random. Next, each couple is assigned a random integer number, t, in {1,2,...,k} which represents the crossover point. The first new string is identical to the original up to position t. The substring from position t onward is swapped with the corresponding substring in the second original string. The second new string is similarly created.

The mutation operator selects a random position of any gene in the population and changes its value from 0 to 1, or viceversa. The probability of this change is approximately one per cent.

4.2 The GA plus the Hopfield net

To mix the neural net and the genetic algorithm, each gene represents an initial point in the neural net. The fitness of the gene is then evaluated by first initialising the neural net, relaxing it to a feasible solution and counting up the assigned frequencies or channels.

The length, k, of the string for a gene has to be equal to the number of neurons in the neural net, that is n times m. The value of a position in the string, 0 or 1, corresponds to the initial point of the associated neuron.

Note that although this mapping constrains the initial points in the net to the vertices of the hypercube (see Section 3.1), no extra benefit can apparently be obtained from a more general initialisation inside the hypercube, as the simulations indicate.

5 SOME SIMULATIONS AND COMPARISONS

We first analyse the performance of the neural network alone for P1 and P2 to provide a frame of reference, and to make a comparison with other networks. Then we apply the hybrid algorithm and compare it to other GAs in literature as well as with the results provided by the net alone. The next section extracts some conclusions from this work and proposes new research directions.

The compatibility matrix for both problems and the demand vector for P2 are given in Table 1. These values correspond to benchmark problem #2 in [Funabiki]. The number of requests for P2 is 167 channels.

2 1 1 0 1 0 1 1 1 1 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0	10
1 2 1 0 1 0 1 1 0 1 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0	11
1 1 2 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0	9
0 0 1 2 0 0 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 1 1 1	5
1 1 1 0 2 0 0 0 0 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0	9
0 0 1 0 0 2 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	4
1 1 1 1 0 1 2 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0	5
1 1 1 1 0 1 1 2 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 1 0	7
1 0 1 1 0 1 1 1 2 1 1 1 1 0 0 0 0 0 0 0 0 0 0 1 1	4
1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 0 0 0 0 0 1 0 1 0	8
0 0 1 1 1 0 1 1 1 1 2 0 1 1 1 1 0 1 1 1 1 1 1 1 1	8
1 1 1 1 1 0 1 1 1 1 0 2 1 1 0 0 0 0 0 0 0 0 0 0 0	9
1 1 1 1 1 0 1 1 0 1 1 1 2 1 1 1 1 1 1 1 0 0 0 0 0	10
1 1 1 0 1 0 0 0 0 1 1 1 1 2 1 1 1 1 1 1 0 0 0 0 0	7
1 1 0 0 1 0 0 0 0 1 1 0 1 1 2 1 1 1 1 1 1 1 0 0 0	7
0 0 0 0 1 0 0 0 0 1 1 0 1 1 1 2 1 1 1 1 0 0 0 0 0	6
0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 2 1 1 0 0 0 0 0 0	4
0 0 0 0 0 0 0 0 0 0 1 0 1 1 1 1 1 2 1 1 0 0 0 0 0	5
0 0 0 0 0 0 0 0 0 0 1 0 1 1 1 1 1 1 2 1 1 1 1 0 0	5
0 0 0 0 0 0 0 0 0 0 1 0 1 1 1 1 0 1 1 2 1 1 1 0 0	7
0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 1 2 1 1 0 0	6
0 0 0 0 0 0 0 0 0 1 1 0 0 0 1 0 0 0 1 1 1 2 1 1 1	4
0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 1 1 1 2 1 1	5
0 0 0 1 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 1 1 2 1	7
0 0 0 1 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 1 1 1 2	5

Compatibility matrix and Requests vector

Table 1.

5.1 The neural net

In order to evaluate the neural net performance, we present a histogram of the solutions it obtains with random initialisation, as well as the histogram of the iterations needed to converge to a feasible solution for P1 and P2.

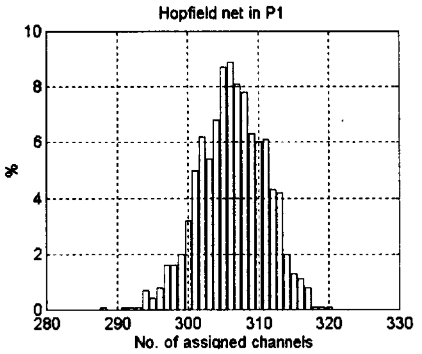


Figure 1.

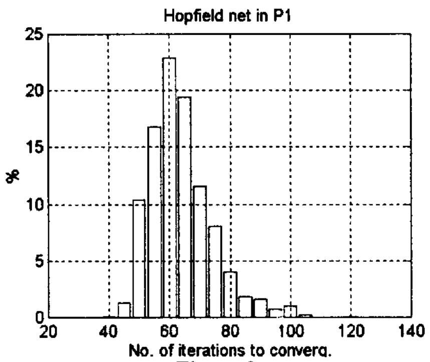


Figure 2.

The introduction of the demand constraints (see Section 3.1) causes a delay in convergence of about 50 iterations with respect to the net for P1. The mean time to convergence for P2 is 110 iterations, but the solutions provided by the network are feasible 100% of the time and just about 2% below the optimum in mean. Note that Funabiki's net gets convergence just 9% of the time and the mean time to converge is 294 iterations.

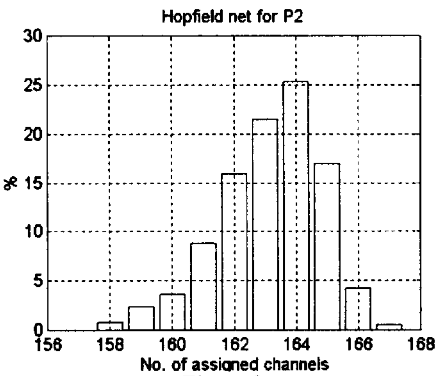


Figure 3.

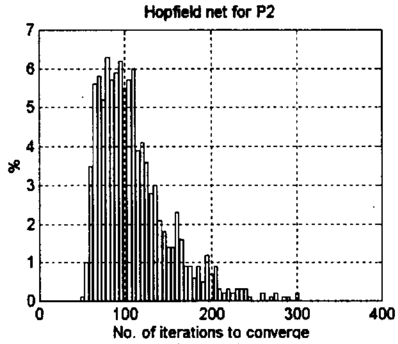


Figure 4.

5.2 The hybrid algorithm

In this section, the hybrid algorithm described in Section 4 is applied to P1 and P2. Results can be observed in Figures 5 and 6, respectively. In each figure, there are given the maximum, the mean and the minimum in the population. The curve at the top is the historic maximum found along the evolution of the genetic algorithm. All these results have been averaged over several realizations.

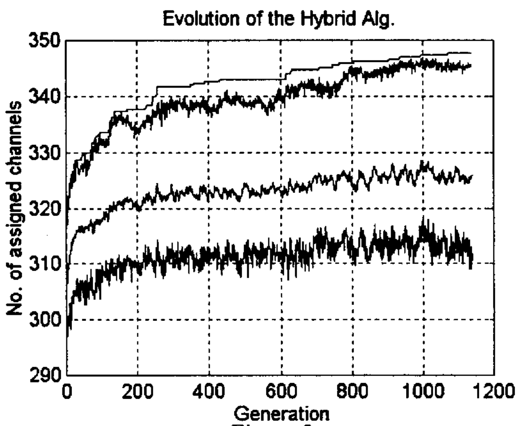


Figure 5.

For P1, the parameters of the GA are those given in Section 4; for P2, since the solutions obtained by the net itself are good enough, we have put a population of just 10 genes. As can be observed in the figures, even with such a reduced population for P2, the optimum is obtained quite frequently through the evolution of the algorithm.

If we compare these results with an standard GA (see for instance [Lai]), we have to observe that the cost function for our algorithm is just the number of assigned channels, while in other GAs this function includes the fitness for the constraints of the problem. Thus, we can guarantee the feasibility of solutions in every generation, and the increase of the cost function directly reflects better assignments.

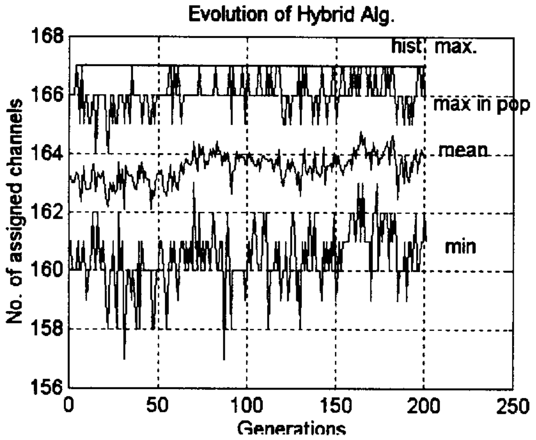


Figure 6.

6 CONCLUSIONS AND FURTHER RESEARCH

In this paper we have proposed a natural manner of merging a local search method based on a Hopfield-type neural network, and a global search method based on genetic algorithms. The design of the neural network allows us to have feasible solutions in every generation of the hybrid algorithm so we can adapt the length of the evolution to the dynamics of the mobile environment. The quality of the allocations obtained by the network is approximately 2% below the optimum for the benchmark problem presented in this paper. For better solutions, the genetic algorithm receives nonperturbed information from the cost function since the constraints are not present in it. For the same reason, the design of genetic operators is simplified.

Since the quality of the solutions provided by the neural network alone may be good enough for some strategies of application, the development of fast genetic algorithms to provide a moderate improvement in the number of assigned channels can suffice for dynamic applications, which have more severe restrictions on speed than on quality. Some other emergent algorithms, such us simulated annealing, can also be combined with neural networks in an analogous manner.

7 REFERENCES

Aarts, E.and Korst, J., "Simulated Annealing and Boltzman Machines", New York, Wiley, 1989.

Bousoño-Calzón, C. and Manning, M., "The Hopfield Neural Network applied to the Quadratic Assignment Problem", Neural Computing and Applications, 3, March 1995, pp 64-72.

Cichocki, A., and Unbehauen, R., Neural Networks for Optimization and Signal Processing, New York: Wiley, 1993

Cox, D.C. and Reudink, D.O., "Layout and Control of high-Capacity Systems" in W.C. Jakes (Ed.), Microwave Mobile Communications, New York: Wiley, 1974.

Funabiki, N. and Takefuji, Y., "A Neural Network Parallel Algorithm for Channel Assignment Problems in Cellular Radio Networks", IEEE Trans. on Veh. Technol., . 41, 4, Nov. 1992, pp.430-437.

Goldberg, D.E., Genetic Algorithms in Search, Optimisation and Machine Learning, Reading, MA: Addison-Wesley, 1989

Hale, W,K., "Frequency Assignment: Theory and Applications", Proc. IEEE, 68, 12, Dec. 1980, pp. 1497-1514.

Hertz, J., Krogh, A., Palmer,R.G., Introduction to the Theory of Neural Computation, Redwood City, CA: Addison-Wesley, 1979.

Hopfield, J.J., "Neural Networks and Physical Systems with Emergent Collective Computational Abilities", Proc. Natl. Acad. Sci. USA, Biophysics, 79, April 1982, pp. 2554-2558.

Kim, S., Kim, S.L., "A Two-Phase Algorithm for Frequency Assignment in Cellular Mobile Systems", IEEE Trans. on Veh. Technol., 43, 3, Aug. 1994, pp. 542-548.

Kunz, D., "Channel Assignment for Cellular Radio Using Neural Networks", IEEE Trans. on Veh. Technol. ,, 40, 1, Feb. 1991, pp. 188-193.

Lai, W.K. and Coghill, G.G., "Channel Assignment Through Evolutionary Optimization", IEEE Trans. on Veh. Technol., 45, 1., Feb. 1996, pp. 91-95.

Lin, F.T., Kao, C.Y., Hsu, C.C., "Applying the Genetic Approach to Simulated Annealing in Solving Some NP-Hard Problems", IEEE Trans. on Systems, Man and Cybernetics, 23, 6, Nov./Dec. 1993 pp. 1752-1767.

Nettleton, R.W. and Schloemer, G.R., " A High Capacity Assignment Method for Cellular Mobile Telephone Systems", Proc. Veh. Technol. Conf., San Francisco, CA, 1989., pp. 359-367.

PAPER TITLE : **Paper 12 Emerging Techniques for Dynamic Frequency Assignment: Merging Genetic Algorithms and Neural Networks**

AUTHOR : C. Bousono-Calzon

NAME : E. Tsang

QUESTION :

1. Can you explain how GA works together with NN?

2. Could you have run GA without NN? Is it true that NN further improves the results in GA?

ANSWER :

1. The GAs genes code points of the unit hypercube. In order to evaluate the fitness of the gene, the NN is initialised with the value coded by the gene and relaxed. The number of assigned channels of the feasible solution given by the network is the fitness value.

2. No. This way of merging both algorithms allows that the information provided to the GA is only that from the objective function, so that, it is not perturbed by constraints. This allows an easier design of GAs and a better performance.

NAME : K.S. Kho

QUESTION :

In the total scope of your work, do you have plans to verify the results of your work in a practical network? For what kind of applications is your algorithm intended to be used?

ANSWER :

Yes. Indeed, I am already making contacts to fund such an application step. The approach presented is quite general. It can be applied to a broad range of assignment problems (with very different types of constraints), to scheduling problems etc.

From a theoretical point of view, it makes it easier to improve the performance of GAs, in that the applications range can be extended by this procedure.

Electromagnetic Coupling Paths to Electronic Systems Connected with Electronic Setups and Destruction Mechanisms

Juergen Bohl, Dr. Tilo Ehlen, Frank Sonnemann
Diehl Foundation & Co., Fischbachstr. 16, D - 90552 Röthenbach

1. Summary

Within an investigation containing both measurements and intensive simulations an effective model creation for the examination of the coupling behaviour of HF interference and the coupling paths into electronic circuits is intended. Based on the predictions deriving from these results adequate hardening measures regarding functionality disturbances due to HF interference can be considered already in the development and design process.

Electromagnetic field simulation programs are necessary for the analysis of the coupling effects into materials, the creation of signal amplifying body resonances and the coupling into power and signal wires. Network analysis programs are responsible for the detection of the HF-LF conversion at the electrical non-linearities and the computation of the radiation of the LF disturbance within the electronic circuit. Figure 1-1 compares the various electromagnetic interferences during the coupling process and the respective simulation programs. A meaningful network analysis requires a coupling of these specific simulation tools.

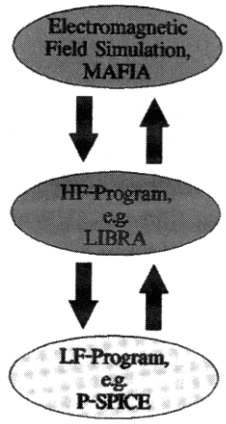


Figure 1-1: Functional Separation of the Simulation Programs Regarding Specific Electromagnetic Interferences

2. The Principle of HF-NF-Conversion

The principle of this HF-LF-conversion or mixture is explained in Figure 2-1. It shows a sinusodial amplitude-modulated HF-interference variable $U_{IN}(t)$ with a carrier frequency of 710 MHz and a modulation frequency of 117 Hz. The frequency domain belonging to this signal lies within the HF-area (Figure 2-2). This HF-interference variable reaches the diode via a wiring (Figure 2-3).

Paper presented at the RTO IST Symposium on "Frequency Assignment, Sharing and Conservation in Systems (Aerospace)", held in Aalborg, Denmark, 5-7 October 1998, and published in RTO MP-13.

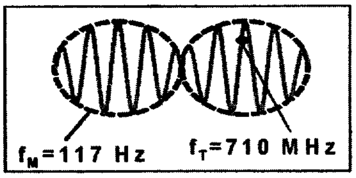


Figure 2-1: Induced HF-Interference $U_{IN}(t)$

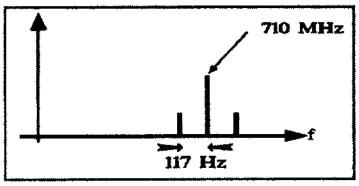


Figure 2-2: Frequency Domain of $U_{IN}(t)$

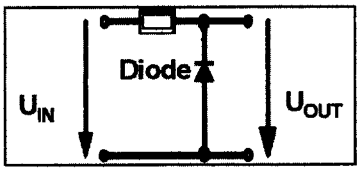


Figure 2-3: Rectifier

For U_{IN} the diode circuit functions as a rectifier. In case of positive half-waves, the diode blocks and the output signal and U_{OUT} follows the input signal. However, in case of negative half-waves, the diode is poled in low resistance direction and U_{OUT} will be short circuited. According to that, the signal after the diode is as shown in Figure 2-4.

Figure 2-5 shows the frequency domain of U_{OUT}. U_{OUT} has a domain within the LF-area in contrary to U_{IN}. It consists of a DC-portion and a multiple of the modulation frequency. A good picture of the domain development is given when you consider U_{OUT} as a scanning signal of its envelope. Within the frequency range of U_{OUT} it is a continuation of its envelope with the scanning frequency.

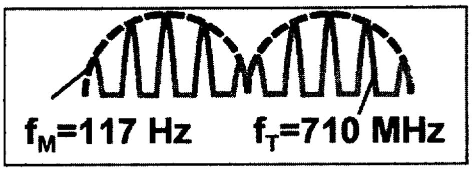


Figure 2-4: $U_{OUT}(t)$ after the Diode

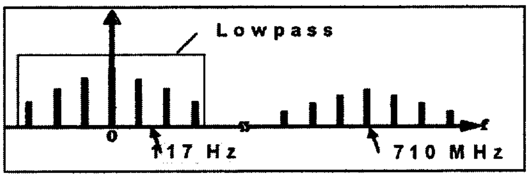


Figure 2-5: Frequency Domain of $U_{OUT}(t)$

The impedance of an electrical NF-circuit is lowpass limited for the HF-range. Any interference of information signals are therefore to be expected only from the LF-part of the domain. Thus, the determined low frequency mixture serves as an input variable for the low frequency circuit analysis.

3. Simulation of HF-Interferences within LF Electronic Circuits

Figure 3-1 shows the electromagnetic circuit coupling. Based on an EM-simulation which calculates the voltages/currents at the end of a wiring the resulting values are the input signals for the electronic circuit which shall contain nonlinear elements. Because of the nonlinear conversion effects, the output contains the LF mix products.

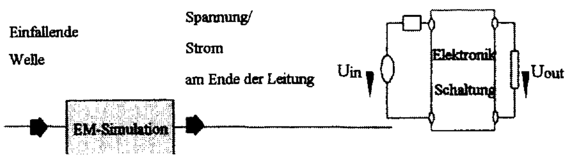


Figure 3-1: Electromagnetic Circuit Coupling

When simulating HF-interferences in LF electronic circuits the calculation of the system response over a broadband frequency range is a problem. In general, this leads to long calculation periods and large data quantities. Althought, the simulation time can be reduced by choosing the respective simulation procedure. In principal, there is a difference between the analytic procedure within the time range and methods within the frequency range. The following comparison of the analytic methods shows a guess of the simulation periods depending on the kind of radiation (CW, pulse).

<u>*Method1 - CW-Time Frame Analysis (PSPICE)*</u>

The following is assumed: a continual, modulated time signal $U_{in}(t)$ with a carrier frequency of 1 GHz and a modulation freuquency of 100 Hz. Due to HF-LF conversion effects (see section 2) low frequency mixture products are created at the output of the electronic circuit. Figure 3-2 shows the input signal. The figure to the right shows the fundamental oscillation (100 Hz) of a LF mixture product.

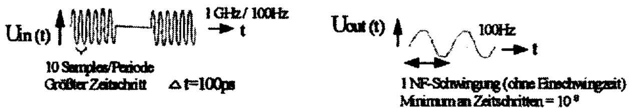


Figure 3-2: HF Input Signal and LF Mixture Product (Time Frame)

At a guess of 10 scanning points/period, the largest time step resulting is a value of $\Delta t = 100$ ps. In order to be able to show only one 100 Hz oscillation the PSPICE simulator has to calculate 10^8 (= 10 ms/Δt) amplitude values.

<u>*Method2 - CW-Frequency Analysis - Harmonic Balance (LIBRA):*</u>

The Harmonic Balance Analysis is based on the fact that after a sufficient period of time a sinusodial oscillation ends in a balanced situation which can be described by a limited Fourier expansion with a respective precision. Figure 3-3 shows the domain $U_{in}(f)$ and the domain $U_{out}(f)$ after a mixture process. Due to the periodicity ot the time signals, the domains consist of single discrete domain lines. To determine a basis of valuation, about 100 - 1000 lines may be sufficient to approximate the signal with sufficient accuracy within the desired frequency range. The number of amplitudes to be calculated per domain line is much lower than the number of calculations at the time frame analysis and therefore the whole procedure is by sizes faster. In case of a CW-oscillation the analyzing method within the frequency range should be preferred against the method within the time frame.

Figure 3-3: HF-Input Signal and Mixture Products at the Exit (Frequency Range)

<u>*Method3-Pulse Time Frame Analysis (PSPICE)*</u>

Figure 3-4 shows a pulse with a carrier frequency of 1 GHz and a pulse width within the µs range. If you guess 10 samples/period, a time step width of $\Delta t = 100$ ps results. Real measurements showed that the response behaviour to the pulse oscillation of the electronics lies within the range of 100 µs to 10 ms. The number of calculations for the unknown amplitudes is within the size of 10^6 to 10^8.

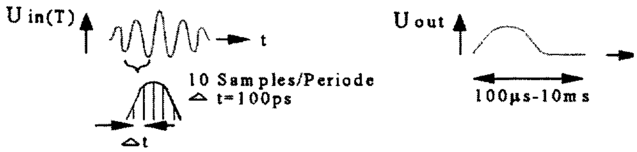


Figure 3-4: HF-Pulse with Response Behaviour of the Electronics (Envelope)

The frequency range method cannot be applied as this procedure assumes a steady-state situation.

With the herein discussed methods for the simulation of HF-interferences into electronic circuits the interferences on the LF information signals can be calculated. Further supportive measurements like the use of system specific simulation tools enable an assessment of the system behaviour leading to a malfunction or not.

For the network simulation it has to be provided that the HF voltage amplitudes do not exceed the boundary values of the active components. In case of an excess of these levels new effects appear within this circuit which were not considered during the modelling. During a very high HF energy supply (as possible for example with an HPM-oscillation) the components can be even destroyed physically. The next paragraph discusses the destroying effects at semiconductors which may occurr during high HPM-oscillation.

4. Destruction Models

The stated destruction models are widely used and follow a plausible physical model assessment although they are not proven in detail. Hereby, you have to differ between primary and secondary destruction effects within the chip. Primary effects are the physical reasons leading to destruction, e.g. electrical voltage flashovers between neighbouring metallized paths on the chip surface. These can lead to a flashover of the isolating layers. Accompanied by a material transport, the flashover leads to a low resistance connection. The reason for this is a very high electrical voltage caused by an HPM oscillation of very high field strength. Further primary effects would be e.g. a junction breakdown or a latch-up effect in case of CMOS components. Secondary effects would be the immediate results leading to a destruction of the component (e.g. metal trace damage, thermal/chemical epoxi reaction). In general, a primary effect leads to a very high voltage flow within the semiconductor. This results in an overheating and the fire loss of conductor paths or substrate material.

4.1 Electrical Flashover (Primary Effect)

The flashover results from fast transient processes with very high voltage between two conductor paths of different potential causing a discharge within the oxide material of the chip surface. Due to the relatively high distance between the conductors, in general there are voltage amplitudes necessary exceeding the breakdown voltages of the semiconductors (pn-junction). Therefore we assume that the transient process is very fast and the time of discharge is finished before the pn-junction becomes conductive. The kind of damage is similar to the ones which can be seen at electrostatic overload (ESD). In this case, the discharge (5 to 25 ns) is done via a small capacitive source (20 - 100 pF) which was charged at a high voltage (2,000 - 20,000 V). Due to the limited energy only a minor damage is done to the oxide layer which can be hardly seen even at high magnification. In case of flashovers with very high energies the following chip surface damages can be detected:

a) physical sites of fracture along the discharge path due to thermo-mechanical load (see Figure 4-1).
b) a local, silvery shining deposit at the border of an AL-conductor path; due to the high current density the Aluminium atoms move along the discharge path and thus generate a visible, cloudy sediment at the silicon oxide (Figure 4-2).

4.2 Junction Breakdown (Primary Effect)

When a pn-junction within a semiconductor is biased backward (non conductive) a charge depleted area is created. In this mode, only very little current (inverse current) flows through the semiconductor. The electrical field accelerates the electrons which themselves partially lose their acquired kinetic energy during their collisions with the atoms of the semiconductor chrystals. If the outer voltage exceeds the breakdown voltage, the kinetic energy is sufficient to open a valence bond in case of a collision with the bar (ionization by collision). A further free electron is created which is also able to generate further electrons in case of a collision with the bar (avalanche effect). The number of the free charge carriers becomes nearly unlimited. The junction becomes low resistant and the immensely increasing current exceeds the normal inverse current by far. As long as the maximum temperature of the junction is not surpassed (e.g. by outer voltage limitation, fitting design of the pn-transition, etc.) the thermal energy of the pn-transition is led to the outside via the chrystal thus not causing any permanent damage. In case of higher energies and in case of exceeding the maximum junction temperature the thermic loading causes a permanent damage to the chrystal structure. Micro fissures and melt spots within the chrystal structure lead to a permanent current leackage overlaying the normal inverse current. The damage of the chrystal is of microscopic extent and embedded below the silicon surface the damage cannot be detected with a standard microscope. Only in case of higher energy impulses the junction breakdown can cause larger damages within the chrystal which can be now detected by microscope (pool of molten metal, cracked dice) and thus leading to a total breakdown of the component.

4.3 Destruction of Metallizing Paths (Secondary Effect)

The current path generated by a flashover and/or the breakdown of a junction is due to an external application (outer wiring) via the connection pin at the semiconductor housing, the bond wire, the Aluminium path at the surface of the semiconductor from the bondpad to the location of the flashover/junction breakdown. Very often the backward path forms within the semiconductor (substrate, bulk) to a GND- or VCC-metallization, then via a bond wire to a connection pin back to the external application. If you consider now the dimensions and the specific conductivity of the stated path elements, the physical and electrical characteristics of the relevant materials and their abilities to dissipate thermal energy to the environment the most endangered elements are the Aluminium paths (< 2μm) on the die surface

Path Element	Material	Profile mm^2	Specific Resistance 10^{-6} Ohm cm	Melting Point °C	Heat Radiation
Conn. Pin	Copper	> 0.1	1.7	1083	not critical
Bond Wiret	Gold/Al	Type $625*10^{-6}$	2.2/2.7	1063/660	not critical
Al-Cond.Path	Aluminium	Type $2*10^{-6}$	2.7	660	bad
Si-Bulk	Silicon	> 0.1		1420	good

Table 4-1: Material Features of the Relevant Path Elements

Figure 4-1: Flashover Bondpad - Ground, Damage to the Oxide Layer by HPM (80 kV/m) Component: 7404AN/TTL (Inverter)

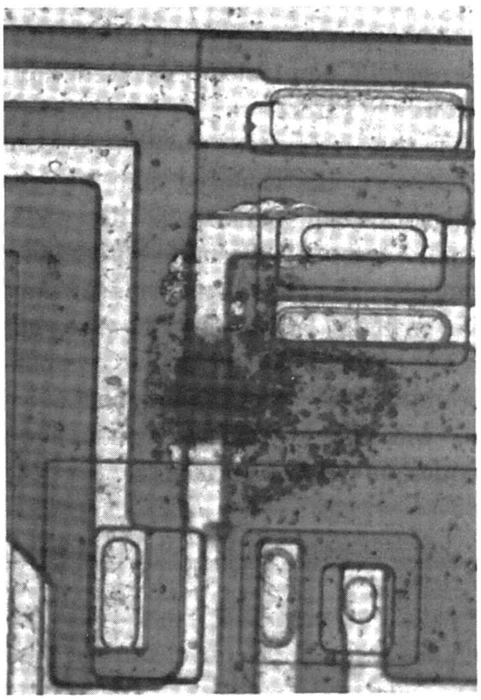


Figure 4-2: Flashover Collector-Emitter Fire Loss of Wiring by HPM (100 kV/m) Component: 7490AN/TTL (Counter)

As the profile of the Aluminium paths on the surface of the chip is at least by the factor 100 smaller (at approximatley the same conductivity) than all other effective path elements, the thermic energy concentrates on the metallizing paths. The thermic energy heats the material of the conductor paths whereas the thinnest points are the most loaded. The passivation layers (protection layers) beneath and above the conductor paths create a thermic barrier so that the thermal energy can hardly be distributed to the environment. The first visible sign of a thermal overcharged metal layer is a decolorization on the Aluminium surface. At a slight increase of the energy the Aluminium melts, contracts and moves away from the hottest spot. The current path is interrupted. The process can be compared with the process of a safety fuse. In case of a high energy transient process, a conductive Aluminium plasma derives from the conductive path material. Due to the high pressure the passivation layer is broken over the conductive path, the plasma explodes and spreads into the surrounding environment.

4.4 Thermical/Chemical Epoxid Reaction (Secondary Effect)

Most of the integrated circuits and other semiconductors are enclosed into an epoxid gum plastics compound. As described in paragraph 4.3, the AL-plasma is spread into the surroundings. This heats up the surrounding package material and melts it. The epoxid decays into several low-molecular components incl. carbon. All these components form together with the Al-plasma a conductive slag sedimenting on the chip surface.

5. Conclusion

The influences of HPM on electronic systems are two-fold: the coupled (HF) interferences and voltages on the wirings caused by HF-LF conversion effects on nonlinearities can cause LF interferences within the electronic wiring overlaying the useful signals and negatively influencing the system performance. With the support of appropriate simulation tools for the circuit analysis the interferences can be numerically calculated. In order to reduce the simulation times as much as possible the methods for analysis shall be dependent on the kind of HF oscillation. The other point is that the HPM can induce such high voltages within the electronic circuit that components are physically destroyed. Although the network analysing programmes allow in general signal levels exceeding the limits of semiconductors, in real life there appear new effects not considered in the modelling of the simulation programs. In no way destruction mechanisms can be rebuilt with the available network tools. Real measurements have to be done for this. A component analysis of destroyed components is used to determine a connection between the HPM radiation characteristics, the component position within the electronic circuit as well as the location and the level of the destruction.

Solving the Radio Link Frequency Assignment Problem using Guided Local Search

Christos Voudouris
Intelligent Systems Research Group
MLB 1/PP 12
BT Laboratories
Ipswich IP53RE
United Kingdom
e-mail: chrisv@info.bt.co.uk

Edward Tsang
Department of Computer Science
University of Essex
United Kingdom
e-mail: edward@essex.ac.uk

1. SUMMARY

In this paper, we examine the application of the combinatorial optimisation technique of Guided Local Search to the Radio Link Frequency Assignment Problem (RLFAP). RLFAP stems from real world situations in military telecommunications and it is known to be an NP-hard problem. Guided Local Search is a metaheuristic that sits on top of local search procedures allowing them to escape from local minima. GLS is shown to be superior to other methods proposed in the literature for the problem, making it the best choice for solving RLFAPs.

2. INTRODUCTION

Guided Local Search (GLS) is a general optimisation technique suitable for a wide range of combinatorial optimisation problems [21]. Successful applications of the technique so far include practical problems such as Workforce Scheduling [20], Frequency Assignment [22] and Vehicle Routing [12]. Classic problems such as the Travelling Salesman Problem (TSP) and the Quadratic Assignment Problem (QAP) have also been tackled with the method. GLS has been shown to be equally good if not better than the best TSP and QAP heuristic search algorithms [23, 21].

GLS was derived from the GENET neural network [25] for Constraint Satisfaction Problems [19] and extends the approach used in GENET to the bulk of combinatorial optimisation problems. GLS belongs to a class of techniques known as *Metaheuristics*. Prominent members of this class include *Tabu Search* [8], *Simulated Annealing* [11], *Genetic Algorithms* [15], and others. Metaheuristics aim at enhancing the performance of heuristic methods in solving large and difficult combinatorial optimisation problems. The method can be used alone or in conjunction with Fast Local Search (FLS). Fast Local Search is a neighbourhood search reduction scheme that particularly suits the operations of GLS [22, 20, 23].

The Radio Link Frequency Assignment Problem (RLFAP) considered in this article is a member of a class of problems known as Frequency Assignment Problems (FAPs). FAPs are NP-hard and they are closely related to the well-known Graph Colouring problem. For a very recent and comprehensive survey of the different FAP variants and problem solving strategies, the reader may refer to [14].

RLFAP stems from real world problems in military telecommunications. A number of RLFAP instances have been released in the framework of the CALMA (Combinatorial ALgorithms for Military Applications) project. The project involved six European research groups and one of its major objectives was to develop and test algorithms for frequency assignment problems. For an overview of the research contacted in the framework of the CALMA project the reader may to refer to [18]. FAPs in general and RLFAP in particular are dealing with the allocation of a limited radio spectrum resource to a number of users in an optimal way.

Guided Local Search and its particular application to the RLFAP described in this paper can be easily extended to other variations of the FAP. In general, the technique is directly applicable to a wider class of problems, known as Partial Constraint Satisfaction Problems [19]. PCSPs extend the classic Constraint Satisfaction Problem (CSP) problem [19] to cover over-constrained instances where no solution exists which satisfies all the constraints. In such cases, we are often seeking solutions which violate the minimum number of constraints.

PCSPs can be used to model a variety of problems where the ever-increasing demand for a limited resource leads to situations where constraints on service quality can be softened. In the context of frequency assignment, constraints that transmitters should not interfere with each other need to be softened because of the limited radio spectrum available, which makes interference unavoidable.

For a detailed description of GLS for PCSPs, they reader may refer to [22, 21]. In this work, we focus on variants of the GLS algorithm for the RLFAP. The paper is structured as follows. We first describe the RLFAP. Following that, we present Local Search, Fast Local Search and Guided Local Search procedures for the problem. The different variants are evaluated on publicly available RLFAP benchmarks and results are compared with a variety of other methods that have been applied to the same instances. The comparisons made clearly indicate the superiority of GLS over alternative approaches proposed for the problem.

Paper presented at the RTO IST Symposium on "Frequency Assignment, Sharing and Conservation in Systems (Aerospace)", held in Aalborg, Denmark, 5-7 October 1998, and published in RTO MP-13.

3. THE RADIO LINK FREQUENCY ASSIGNMENT PROBLEM

The *Radio Link Frequency Assignment Problem* (RLFAP) was abstracted from the real life application of assigning frequencies to radio links. The interference level between the frequencies assigned to the different links has to be acceptable; otherwise communication will be distorted. The frequency assignments have to comply with certain regulations and physical characteristics of the transmitters. Moreover, the number of frequencies is to be minimised, because each frequency used in the network has to be reserved at a certain cost. In certain cases, some of the links may have pre-assigned frequencies which may be respected or preferred by the frequency assignment algorithm. RLFAP is NP-Hard and it is a variant of the graph T-colouring problem as introduced by Hale [9]. IP formulations of the problem can be found in [14]. Following [18], the problem can be briefly described as follows.

We are given a set L of links. For each link i a frequency f_i has to be chosen from a given domain D_i. Some links may already have a pre-assigned frequency pf_i, which may or may not be changed. Constraints are defined on pairs of links that limit the choice of frequencies for these pairs. For a pair of links $\{ i, j \}$ these constraints are either of type

$$|f_i - f_j| > d_{ij} \qquad (1)$$

or of type

$$|f_i - f_j| = d_{ij} \qquad (2)$$

for a given distance $d_{ij} \geq 0$. Two links i and j involved in a constraint of type (1) are called *interfering* links, and the corresponding d_{ij} is the interfering distance. Two links bound by a constraint of type (2) are referred to as a pair of *parallel* links; every link belongs to exactly one such pair.

Some of the constraints may be violated at a certain cost. Such restrictions are called *soft*, in contrast to the hard *constraints*, which may not be violated. The constraints of type (2) are always hard, and so are some of the pre-assignment constraints which dictate the preferred frequency for a link also called *mobility constraints*. Interference costs c_{ij} for violating soft constraints of type (1) and mobility costs m_i for changing soft pre-assigned frequencies are given. An assignment of frequencies is complete if every link in L has a frequency assigned to it. We denote by C and M the sets of all soft *interference* and *mobility* constraints, respectively.

The first priority is to find a complete assignment that satisfies all hard constraints and is of minimum cost:

$$\min \sum_{C} c_{ij} \cdot \delta(|f_i - f_j| \leq dij) + \sum_{M} c_i \cdot \delta(|f_i - pf_i| > 0) \qquad (3)$$

subject to hard constraints:

$|f_i - f_j| > d_{ij}$: for all pairs of links $\{ i, j \}$ involved in the hard constraints,

$|f_i - f_j| = d_{ij}$: for all pairs of parallel links $\{ i, j \}$,

$f_i = p_i$: for all links $i \in$ L with hard pre-assigned frequencies,

$f_i \in D_i$: for all links $i \in$ L,

where $\delta(.)$ is 1 if the condition within brackets is true and 0 otherwise.

If there exists a feasible assignment (i.e. a complete assignment of zero cost) then we would like to find a feasible assignment that satisfies all constraints and minimises the number of distinct frequencies used or the largest frequency used:

$$\min |\cup_i \{ f_i \}| \qquad (4)$$

or

$$\min \max_i f_i \qquad (5)$$

subject to the hard and soft constraints:

$|f_i - f_j| > d_{ij}$: for all pairs of links $\{ i, j \}$ involved in the soft and hard constraints,

$|f_i - f_j| = d_{ij}$: for all pairs of parallel links $\{ i, j \}$,

$f_i = p_i$: for all links $i \in$ L with hard pre-assigned frequencies,

$f_i \in D_i$: for all links $i \in$ L.

Essentially, there are two distinct types of RLFAP instances. Instances where all solutions violate one or more constraints of the problem. For these *insoluble* instances, the cost function is given by (3). Instances for which exist at least one solution which satisfies all the constraints. For these *soluble* instances, we are seeking solutions which satisfy the constraints and either minimise (4) or (5). In the following, we describe the proposed GLS algorithm for the problem. We begin with local search which is the basis of the method.

4. LOCAL SEARCH FOR THE RLFAP

A local search procedure for PCSPs which is also applicable to the RLFAP has been described in previous work of the authors [22]. The scheme is based on the min-conflicts heuristic of Minton et al. [13] for Constraint Satisfaction Problems and also the computational model of the GENET neural network [25, 6]. An 1-optimal type move is used which changes the value of one variable at a time. Starting from a random and complete assignment of values to variables, variables are examined in an arbitrary static order. Each time a variable is examined, the current value of the variable changes to the value (in the variable's domain) which yields the minimum value for the objective function. Ties are randomly resolved allowing moves which transit to solutions with equal cost. These moves are called *sideways moves* [17] and enable local search to examine plateau of solutions occurring in the landscapes of many SAT and CSP problems.

The above local search procedure can be directly applied to RLFAP by considering each link as a variable with a domain given by all the possible frequencies that can be assigned to the link. Results on using this approach were reported in [22]. A more efficient approach, used in this work, takes into account the fact that each link in RLFAP is connected to exactly one other link via a hard constraint of type (2). In particular, we can define a local search variable for each pair of parallel links bound by an equality constraint [7]. The domain of this variable is defined as the set of all pairs of frequencies from the original domains of the parallel links which satisfy the hard equality constraint. Next, we examine how this basic local search procedure can be improved by taking advantage of the Fast Local Search approach.

5. FAST LOCAL SEARCH FOR THE RLFAP

Best improvement local search for the RLFAP as used in the context of Tabu Search [1, 7] evaluates all possible 1-optimal moves over all variables before selecting and performing the best move. Given the large number of links in real world instances, greedy local search is a computationally expensive option. The local search procedure described in previous section is already a faster alternative. An even faster alternative is fast local search initially used in the TSP under the name *don't look bits* technique for solving large instances of the problem [2]. Fast local search has been abstracted by Voudouris and Tsang and used in a number of applications as a generic neighbourhood reduction scheme [22, 20, 23].

In the case of the RLFAP, a bit is attached to each variable. If the bit of the variable is 1 then the variable is called *active* and it is examined for improving moves otherwise it is called *inactive* and it is discarded by local search. Whenever a variable is examined and its value is changed (i.e. the variable's parallel links are assigned to another pair of frequencies because of an improving or sideways move) the activation bit of the variable remains to 1 otherwise it turns to 0 and the variable is excluded in future iterations of the improvement loop. Additionally, if a move is performed activation spreads to other variables which have their bits set to 1. In particular, we set to 1 the bit of variables where improving moves may occur as a result of the move just performed. These are the variables for which either one of their links is connected via a constraint to one of the links of the variable that changed value. There are five potential schemes for propagating activation after changing the value of a variable. They are the following:

S1. Activate all variables connected via a constraint to the variable which changed value.

S2. Activate only variables that are connected via a constraint which is violated. This resembles CSP local search methods where only variables in conflict have their neighborhood searched.

S3. Activate only variables that are connected via a constraint which has become violated as a result of the move (subset of S2 and also S4).

S4. Activate only variables that are connected via a constraint that changed state (i.e. violated → satisfied or satisfied → violated) as a result of the move (superset of S3).

S5. Activate variables that fall under either condition S2 or S4.

Fast local search stops when all the variables are inactive or when a local minimum is detected by other means. In the following, we describe the metaheuristic technique of GLS which significantly improves the performance of local searches.

6. GUIDED LOCAL SEARCH

Guided local search is a metaheuristic algorithm which its aim is primarily to help local search to escape local optima. In doing so, the algorithm is also trying to distribute the search efforts to the best effect by guiding local search towards solutions of high quality. The basic idea is to augment the objective function with penalties, which direct the search away from local optima and towards high quality solutions.

To apply GLS, one has to define *features* for the candidate solutions. For example, in the travelling salesman problem, a feature could be "whether the candidate tour travels immediately from city A to city B" [23]. GLS associates *costs* and *penalty* to each feature. The costs should normally take their values from the objective function. For example, in the travelling salesman problem, the cost of the above feature is the distance between cities A and B. The penalties are initialised to 0 and will only be increased when the local search reaches local optimum. This will be elaborated below.

Given an objective function g that maps every candidate solution s to a numerical value, we define a function h which will be used by local search (replacing g).

$$h(s) = g(s) + \lambda \times \Sigma(p_i \times I_i(s)) \quad (6)$$

where s is a candidate solution, λ is a parameter to the GLS algorithm, i ranges over the features, p_i is the penalty for feature i (all p_i's are initialised to 0) and I_i is an indication of whether s exhibits feature i:

$$I_i(s) = 1 \text{ if } s \text{ exhibits feature } i\text{; 0 otherwise} \quad (7)$$

When local search settles on a local optimum, the penalty of some of the features associated to this local optimum is increased (to be explained below). This has the effect of changing the objective function (which defines the "*landscape*" of the local search) and driving the search towards other candidate solutions. The key to the effectiveness of GLS is in the way that penalties are imposed. It is worth pointing out that a slight variation in the way that penalties are managed could make all the difference to the effectiveness of a local search.

Our intention is to penalise "bad features", or features which "matter most", when a local search settles in a local optima. The feature that has high cost affects the overall cost more. Another factor that should be considered is the current penalty value of that feature. We define the utility of penalising feature *i*, *utili*, under a local optimum s_*, as follows:

$$Util_i(s_*) = I_i(s_*) \times c_i / (1 + p_i) \quad (8)$$

where c_i is the cost and p_i is the current penalty value of feature i. In other words, if a feature is not exhibited in the local optimum, then the utility of penalizing it is 0. The higher the cost of this feature (c_i), the greater the utility of penalising it. Besides, the more times that it has been penalised, the lower the utility of penalising it again.

In a local optimum, the feature(s) with the greatest *util* value will be *penalised*. This is done by incrementing its penalty value by 1:

$$p_i = p_i + 1 \quad (9)$$

By taking cost and the current penalty into consideration in selecting the feature to penalise, we are distributing the search effort in the search space. Candidate solutions which exhibit "good features", i.e. features involving lower cost, will be given more effort in the search, but penalties help to prevent all effort be directed to the best features. Following we shall describe the general GLS procedure:

Procedure ***GLS*** *(input: an objective function g; a local search strategy L; features and their costs; parameter λ)*

0. Generate a starting candidate solution randomly or heuristically;

1. Initialize all the penalty values (p_i) to 0;

2. Repeat the following until a termination condition (e.g. a *maximum number of iterations or time limit) has been reached:*

*3.1. Perform local search (using **L**) according to the function h (which is g plus the penalty values, as defined in (6) above) until a local optimum M has been reached;*

3.2. For each feature i which is exhibited in M compute $util_i = c_i / (1 + p_i)$

3.3. Penalize every feature i such that $util_i$ *is maximum:* $p_i = p_i + 1$;

Return the best candidate solution found so far according to the objective function g.

Applying guided local search to a problem requires the existence of a local search procedure preferably a version of fast local search and also a set of features which will be used to guide local search in the problem's search space. If a fast local search variant is available, we can combine it with GLS in a straightforward way. The key idea is to associate features to activation bits. The associations to be made are such that for each feature we know which variables contain moves that have an immediate effect upon the state of the feature (i.e. moves that remove the feature from the solution).

At the beginning of GLS, all the activation bits of fast local search are set to 1 and fast local search is left to reach the first local minimum. Whenever a feature is penalised, the bits of the associated variables and only these are set to 1. In this way, after the first local minimum, fast local search calls start by examining a subset of the variables and in particular those which associate to the features just penalised.

Another variant to the GLS method which seems to significantly improve performance in certain RLFAP instances is to decrease penalties and not only increase them. More specifically, the variant uses a circular list to retract the effects of penalty increases made earlier in the search process, in a way that very much resembles a tabu list. In particular, penalties increased are decreased after a certain number of penalty increases is performed. The scheme uses an array of size t where the t most recent features penalised are recorded. The array is treated as a circular list, adding elements in sequence in positions 1 through t and then starting over at position 1. Each time the penalty of a feature is increased (by one unit), the feature is inserted in the array and the penalty of the feature previously stored in the same position is decreased (by one unit). The rational behind the strategy is to allow GLS to return to regions of the search visited earlier in the search process, so introducing a search intensification mechanism.

Next, we examine the objective function g used by local search in the RLFAP, the GLS features defined, their costs and also how the problem variables are activated in the case of fast local search when these features are penalised.

7. LOCAL SEARCH AND GLS APPLIED TO THE RLFAP

7.1 Objective Function

In the RLFAP, we defined and used a simple objective function for both insoluble and soluble instances. The objective function g was given by the sum of all constraint violation costs in the solution with all the constraints contributing equally to the sum. This objective function is as follows:

$$g(s) = \sum_{C \cup C^{Hard}} \delta(|f_i(s) - f_j(s)| \leq dij) + \sum_{M} \delta(|f_i(s) - pf_i| > 0) \qquad (10)$$

subject to hard constraints:

$f_i(s) \in D_i'$: for all links $i \in$ L,

where $\delta(.)$ is 1 if the condition within brackets is true and 0 otherwise, $f_i(s)$ is the frequency assigned to link i in solution s, C^{Hard} is the set of hard inequality constraints, C is the set of soft inequality constraints, M is the set of soft mobility constraints and D_i' is the reduced domain for link i containing only frequencies which satisfy the hard equality and mobility constraints. A solution s with cost 0 with respect to g is satisfying all hard and soft constraints of the problem.

GLS allows us to use such a simple objective function for local search and still obtain high quality solutions. Most of the costs contained in the original cost function of the problem can be minimised by defining features for solutions and setting the feature costs to appropriate values. The application of penalties can force local search toward solutions of high quality, to some degree independently from the objective function used by local search.

GLS will normally perform better with an objective function which is closely based on the cost function of the problem. The motivation to use a simple function such as (10) is closely related to the rugged landscapes formed in RLFAP, if the original cost function is used. In particular, high and very low violation costs are defined for some of the soft constraints in insoluble instances. This leads to even higher violation costs to have to be defined for hard constraints. The landscape is not smooth but full of deep local minima mainly due to the hard and soft constraints of high cost. Soft constraints of low cost are buried under these high costs. CALMA researchers, implementing Simulated Annealing for the insoluble instances, also found it difficult to devise a cooling regime which is effective on these types of landscapes [7].

In the following, we continue by presenting the different features defined in the RLFAP to help local search escape local minima and also be guided towards solutions of high quality.

7.2 Constraint Features for Minimising Interference Costs

One very important cost factor in the RLFAP is the constraint violation costs defined for soft inequality constraints. Inequality constraints can be used to define a basic feature set for the RLFAP. Each inequality constraint is interpreted as a feature with the feature cost given by the violation cost of the constraint c_{ij} as defined in the problem's original cost function (3).

For the RLFAP, the objective function of local search is augmented with a set of modifiable penalty parameters one for each inequality constraint. Initially, the penalty parameters are set to 0 and each constraint accounts only for its violation cost set to 1 for all constraints. Each time local search settles in a local minimum, the penalties for some of the constraints violated (the corresponding features are exhibited) are increased according to the general scheme described in section 6. Constraints with high cost are penalised more frequently than those with low cost. In the short term, local search escapes from the local minimum while in the long term, it is biased to spend more time on solutions that satisfy high cost constraints than low cost ones.

Hard inequality constraints are also modelled as features though the cost assigned to them is infinity. This results in their utility to be penalised to also tend to infinity. To implement that in the program, hard constraints are given priority over all other features (i.e. features defined for soft inequality constraints and other features to be introduced below). This basically forces local search to return back to a feasible region where penalising other features can resume.

7.3 Value Features for Minimising Mobility Costs

The mobility constraints considered during the search process are all of soft nature since the hard mobility constraints are pre-processed. To minimise this cost factor, a feature is defined for each value in the domain of a variable and that for all variables which have mobility constraints associated with their links. The cost of each *value feature* is defined by the sum of mobility costs m_i that would be incurred according to (3) if the pair of frequencies represented by the value were assigned to the links modelled by the variable.

Each time local search settles in a local minimum, values assigned to variables with mobility constraints or constraints violated (i.e. the corresponding features are exhibited) will have their penalties increased according to the general scheme described in section 6. If GLS is combined with the fast local search of section 5 then variables associated with the penalised constraints or values will also be activated to focus fast local search to remove the penalised features.

7.4 Frequency Features for Minimising the Number of the Frequencies Used

Features for minimising this factor are only deployed in soluble instances. In particular, a *frequency feature* is defined for each frequency appearing in the union of domains of the links. In contrast to other features described above, the penalty for a frequency feature is incorporated in the objective function not once but multiple times. More specifically, for each variable that can use a specific frequency, the penalty for this frequency is incorporated in the objective function multiplied by an indication function. This indication function takes the value 1, if the frequency is utilised by the variable in at least one of its parallel links or 0 otherwise. This enable us to force all variables that are using a specific frequency to avoid it by simply increasing the penalty for that frequency. Given the capability to force local search to avoid specific frequencies, the question is how to choose these frequencies so that the overall number of frequencies used is minimised during the search process.

A simple heuristic is to increase the penalties for frequencies used by only few of the variables. Eliminating these frequencies will eventually lead to a decrease in the total number of frequencies used since variables will be in their majority assigned to frequencies used by many variables. More formally, the feature cost for a feature representing frequency $f \in \cup_i \{ f_i \}$ is defined as follows:

$$c_f(s) = 1/|\,L_f(s)| \qquad (11)$$

where L_f is the set of links assigned to the frequency f in solution s. Obviously, frequencies not assigned to any of the links in the local minimum may be discarded since penalising them will have no effect. The indication function for a frequency feature to be incorporated in its utility function is as follows:

$$I_f(s_*) = 1 \text{ if } s_* \text{ uses frequency } f; \text{ 0 otherwise.} \qquad (12)$$

Frequency features maximising the utility function (8) (as it is instantiated by using (11) and (12)) have their penalties incremented.

If GLS is combined with fast local search, we activate all variables which are using the penalised frequency in either one of their links. Finally, frequency features are only considered for penalisation if no soft or hard constraints are violated since we want to minimise these cost factors first.

7.5 Value Features for Minimising the Maximum Frequency Used

For minimising this optimisation criterion, we introduced a feature set similar to that used for mobility costs. As with mobility costs, one feature was defined for each value in the domain of a variable though this time the feature costs were all equal and set to the value 1. For the penalties on values to have an effect in minimising the maximum frequency used, we only considered for penalising values which were responsible for assigning the maximum frequency used in the local minimum. Since all features are having an equal cost, the utility function of GLS simply translates in selecting the feature(s) which has been penalised the minimum number of times. As before, if fast local search is utilised, the variable of a value penalised is activated.

8. EXPERIMENTAL EVALUATION OF GLS

8.1 Experimental Setting

We conducted a large number of experiments on the publicly available RLFAP instances. For all GLS variants tested, the λ parameter was set to 1 and ten runs were performed from different random initial solutions for each benchmark instance. Since GLS is not a complete method, a termination criterion has to be specified. For that purpose, we set a limit on the number of 1-optimal moves that are evaluated in a single run. The value chosen for that limit was $4\text{x}10^7$ moves for all variants and instances considered.

In order to give an indication of the times spent on problem solving, we also measured the *User Time* (excluding any *System Time* such as i/o operations, memory allocation etc.) required by the variants to reach the best solution in each of their runs. All times reported refer to a SPARCUltra workstation running at 168Mhz which represents a typical modern UNIX workstation.

To evaluate the performance of variants with regard to the quality of solutions they can produce, we used the measure of percentage **excess** above the best known solution. This measure is calculated in the following way:

excess = 100 * (solution_cost – best_known_solution_cost) / best_known_solution_cost

In total, we used 25 instances to test GLS and its variants. Eleven of these instances were made publicly available by the French Centre Electronique de l' Armament. These instances are based on real world problems and are considered to be representative of the different variations of the problem. We will refer to these instances as the *Celar* set.

In addition to these instances, a set fourteen were made available during the CALMA project. These instances were generated by the GRAPH problem generator described in [3].

Although random in nature, they thought to resemble the realistic Celar instances. We will refer to this second set of instances as the *Graph* set. In the following, we highlight some of the results from the experiments conducted. A detailed report on the results and the also the GLS method on the RLFAP will be provided in [24].

8.2 Guided Local Search Variants for the RLFAP.

In total, there are six variations of GLS for the RLFAP problem. The variants are derived by combining GLS with the local procedures presented in sections 4 and 5. These variants are the following:

- GLS-FI: This is the result of the combination of GLS with first improvement local search of section 4.
- GLS-FLS-S1, GLS-FLS-S2, GLS-FLS-S3, GLS-FLS-S4, GLS-FLS-S5: This is the set of variants resulting from the combination of GLS with fast local search using one of the available activation schemes (see section 5).

We tested all the above techniques on the RLFAP instances and also experimented with the circular list strategy when combined with GLS-FLS-S3, GLS-FLS-S4, and GLS-FLS-S5.

8.3 Best Known Solutions and GLS

GLS variants repeatedly found or improved the best known solutions as reported in [18] and also in an Internet Web page by Thomas Schiex devoted to the RLFAP instances (http://www-bia.inra.fr/T/schiex/Doc/CELARE.html). Table 1 summarises this information. In this table, solution costs in italic characters indicate that GLS was able to find the best known solution, while in bold indicate that GLS improved the best known solution. An asterisk next to a solution cost indicates that the cost is known to be optimal.

For the 25 publicly available RLFAP instances, GLS managed to find the best known solution in 18 of them, improved the best known solution in 5 and found a marginally inferior solution in only 2 of the instances. Note here, that the best known solutions have been discovered by different techniques. GLS achieved a better performance than the collection of these techniques by improving the cost of the previously best known solutions by an average 1,75%.

8.4 Comparing the GLS Variants.

All the GLS variants achieved very good results. It is worth pointing that despite the thousands of runs performed for testing the 6 variants, GLS was always able to find a feasible solution for the soluble instances and that despite the fact that hard inequality constraints were included in the cost function.

As expected, GLS-FLS variants were better on average than the GLS-FI variant. GLS-FLS-S5 was found to be the best variant with regard to the measures of average excess and worst excess. We also found that all GLS-FLS variants with the exception of GLS-FLS-S1 are well suited for finding high quality solutions (i.e. they have an average *best excess* between 1% and 2% above the best known solutions). GLS-FLS-S1 is activating more variables that the other FLS variants, this seems to be having an adverse effect in computation times which in each turn affects the quality of solutions found by GLS-FLS-S1. The variants considered in this section are not using the circular list strategy for penalties.

As we are going to see in the next section , this strategy can significantly improve the GLS performance.

RLFAP Instance	Best Known Solution Cost	Best Solution Cost found by GLS
celar1	16*	*16*
celar2	14*	*14*
celar3	14*	*14*
celar4	46*	*46*
celar5	792*	*792*
celar6	3389*	*3389*
celar7	343592	343598
celar8	262	*262*
celar9	15571	*15571*
celar10	31516	*31516*
celar11	22*	24
graph1	18*	*18*
graph2	14*	*14*
graph3	380*	*380*
graph4	394*	*394*
graph5	221	*221*
graph6	4189	**4123**
graph7	4324	*4324*
graph8	20	**18**
graph9	18*	*18*
graph10	394*	*394*
graph11	3513	**3081**
graph12	11827	*11827*
graph13	11130	**10119**
graph14	10	**8***

Table 1 Comparison between the best known solutions and the best solutions found by GLS.

8.5 Results Using the Circular List Strategy

GLS variants can be further improved by utilising the circular list strategy described in section 6. We selected the variants GLS-FLS-S5, GLS-FLS-S4 and GLS-FLS-S3 to experiment with this strategy. GLS-FLS-S4 produced the best results when combined with the strategy while GLS-FLS-S5 and GLS-FLS-S3 also improved their performance. Results for GLS-FLS-S4 using the circular list strategy are reported in Table 2. As we can see in this table, GLS-FLS-S4 achieves a stunning *average excess* of only 1.63% above the best known solutions. This is a massive improvement over the best *average excess* of 8,83% for GLS variants not using the circular list, achieved by GLS-FLS-S5 in the experiments of section 8.4.

Another result worth noting in Table 2 is that the best known solutions were found in each and every run of GLS-FLS-S4 for eleven of the instances (see rows with mean excess equal to 0). The average user time for the variant was a reasonable 4.73 minutes. Of course when used in practice, the algorithm can be stopped at any time and report the best solution found up to that point.

RLFAP Instance	Best Solution Cost	Worst Solution Cost	Mean Excess (%)	Mean User Time (min)
celar1	16	18	3.75	5.63
celar2	14	14	0	0.02
celar3	14	14	0	0.19
celar4	46	46	0	0.01
celar5	792	792	0	0.09
celar6	3389	3415	0.312777	15.16
celar7	343702	354205	0.442181	17.55
celar8	262	285	3.396947	15.68
celar9	15571	15612	0.04046	6.62
celar10	31516	31517	0.000635	5.46
celar11	24	30	22.72727	11.55
graph1	18	18	0	1.70
graph2	14	14	0	0.04
graph3	380	380	0	0.08
graph4	394	394	0	0.37
graph5	221	221	0	0.48
graph6	4133	4167	0.603929	11.84
graph7	4329	4338	0.157262	6.77
graph8	18	18	0	4.33
graph9	18	20	1.111111	1.67
graph10	394	764	9.390863	0.59
graph11	3087	3137	1.194417	4.67
graph12	11832	11840	0.049886	3.84
graph13	10194	11334	5.691274	3.92
graph14	8	8	0	0.08
Mean			1.627129	4.73

Table 2 Results for the variant GLS-FLS-S4 using the circular list strategy.

8.6 Comparison with the CALMA Project Algorithms.

In the framework of the CALMA project, a number of optimisation techniques were applied to the RFLAP, amongst them versions of Simulated Annealing, Tabu Search and Genetic Algorithms. A summary of the results for the techniques tested on the RLFAP by the CALMA project members is given in [18]. To give an indication of the relative performance of GLS with regard to these techniques, we compare the results reported in [18] with the results obtained by GLS-FLS-S4 using the circular list strategy. Tables 3 & 4 compare GLS results with the results from the CALMA techniques on the Celar and Graph sets respectively.

A number of conclusions can be drawn based on the information presented in these tables. Firstly, GLS has been successfully applied to all instances without exception something which is not the case for any of the other algorithms. More importantly the performance of GLS is consistent across all instances. The only techniques which offer a wide coverage of instances and produce results consistent and competitive with those of GLS are the Simulated Annealing variant by EUT [7], the Variable Depth Search variant by EUT [7], and the Extended GENET by KCL which is a neural network [16, 1].

The results for Extended GENET are exceptionally good for a neural network architecture applied to a CO problem. Nonetheless, they are not surprising since the technique is also based on the GENET neural network architecture like GLS. The interested reader may refer to [22] for a direct comparison between Extended GENET and the less efficient version of GLS presented in there. In brief, we found in [22] that the average solution quality produced by Extended GENET is much inferior than that of GLS as implemented in that paper. The main reason being that Extended GENET does not use some of the principles of GLS such as selective penalisation and problem modelling using features. Furthermore, Extended GENET does not apply to problems with mobility costs and it is cumbersome to extend since a problem definition has to be cast to a neural network architecture.

The Simulated Annealing variant by EUT is performing well on both the soluble and insoluble instances of the CELAR set though overall the results are inferior in terms of solution quality than those of GLS. Furthermore the algorithm is very time intensive on the insoluble instances. This last behaviour may be related to the fact that it is difficult to find a good cooling regime for the rugged landscapes of RLFAP's insoluble instances as this is pointed out in [7].

In our opinion, the Variable Depth Search by EUT is the most competitive algorithm to the GLS-FLS-S4 variant though on average has an inferior performance. In particular, GLS has a significantly better performance on the insoluble instances of the Graph set and also the soluble instances of the CELAR set. A better solution by VDS is only reported for instances Celar7 and Graph7 and that by a very small margin. Furthermore the technique is not applied to instances were the maximum frequency is minimised.

Generally speaking, if the average behaviour of the CALMA methods was well documented then further conclusions could be drawn. As mentioned above, GLS-FLS-S4 has a very small variance with regard to solution quality (see Table 2) and that could be a significant advantage over other heuristic methods which in many cases, they are very sensitive to the random starting point chosen in each run.

Finally, Tabu Searches (TS) developed by the CALMA project members are not very competitive, the main reason being the use of best improvement local search which given the large number of variables is a very expensive option for these problems. Further work on Tabu Search variants using candidate list strategies and/or move update schemes may result in better performance for Tabu Searches in the RLFAP. Successful TS variants have already been developed though in the context of the relatively simpler minimum interference problem [4, 5, 10]. RLFAP introduces a number of additional requirements such as mobility constraints, hard inequality and equality constraints, soft constraints with different violation costs, minimisation of the number of frequencies used, minimisation of the maximum frequency used, and also different domains for the problem's variables. At the moment, efficient TS approaches such as those described in [10] are only addressing the minimisation of constraints with equal violation costs and to some extent the problem of minimising the number of frequencies used, though all variables are consider to have the same domain and the minimisation of the criterion is conducted in a rather ad-hoc way (i.e. multiple runs with fixed frequency sets [10]).

9. CONCLUSIONS

In this paper, the application of Guided Local Search to the Radio Link Frequency Assignment Problem was examined. A number of different variants were described and evaluated on publicly available RLFAP benchmarks instances. We found that the GLS method is an ideal technique for the problem finding solutions of high quality in reasonable time while being applicable to all different variations of the problem. In addition to that, comparisons of GLS with other methods demonstrated the superiority of the technique.

Given the very good performance of Guided Local Search on the problem (i.e. the best GLS variant achieves an average excess of 1,63% over the best known solutions for the benchmark instances), we believe there is little room for improvement with regard to solution quality by using more sophisticated techniques. Future research, related to this work, will focus on the application of GLS in practical systems for frequency assignment and other problems. The real world is the ultimate testbed for evaluating an algorithm given the many difficulties involved such as side constraints, complex objective functions, user preferences, and also requirements for dynamic or distributed optimisation.

REFERENCES

[1] Boyce, J.F, Dimitropoulos, C.H.D., vom Scheidt, G., Taylor, J.G., "GENET and Tabu Search for combinatorial optimisation problems", *Proceedings of World Congress on Neural Networks*, Washington D.C., INNS Press, 1995.

[2] Bentley, J.L., "Fast Algorithms for Geometric Traveling Salesman Problems", *ORSA Journal on Computing*, Vol. 4, 387-411, 1992.

[3] Benthem, H., Hipolito, A., Jansen, B., Roos, C., Terlaky, T., and Warners, J., "GRAPH: a test case generator, generating radiolink frequency assignment problems heuristically", *EUCLID CALMA technical report*, EUCLID CALMA Project, Delft and Eindhoven Universities of Technology, The Netherlands, 1995.

[4] Castelino, D.J., Hurley, S., and Stephens, N.M., "A Tabu search algorithm for frequency assignment", *Annals of Operations Research*, Vol 41, pp. 343-358, 1993.

[5] Castelino, D., and Stephens, N., "Tabu thresholding for the frequency assignment problem", *Meta-Heuristics: Theory and Applications* (Edited by I.H. Osman and J.P. Kelly), Kluwer Academic Publishers 1996.

[6] Davenport, A., "GENET: a connectionist architecture for constraint satisfaction", *PhD Thesis*, Department of Computer Science, University of Essex, 1997.

[7] Eindhoven RLFAP Group, "EUCLID CALMA, Radio Link Freuquency Assignment Project, Technical Annex-2.3.3: Local search", *EUCLID CALMA technical report*, EUCLID CALMA Project, Eindhoven University of Technology, The Netherlands, 1995.

[8] Glover, F., and Laguna, M., *Tabu Search*, Kluwer Academic Publishers, Boston, 1997.

[9] Hale, W.K., "Frequency Assignement: Theory and Applications", *Proceedings of the IEEE*, 68, 1497-1514, 1980.

[10] Hao J.-K., Dorne, R., Galinier, P., "Tabu Search for Frequency Assignment in Mobile Radio Networks", *Journal of Heuristics*, Volume 4, Issue 1, pp. 47-62, June 1998.

[11] Kirkpatrick, S., Gelatt, C.D., and Vecchi, M.P., "Optimisation by Simulated Annealing", *Science*, Vol. 220, 671-680, 1983.

[12] Kilby, P., Prosser, P., and Shaw, P., "Guided local search for the vehicle routing problem", *Proceedings of the 2nd International Conference on Metaheuristics*, July 1997.

[13] Minton S., Johnston, M.D., Philips A.B., and Laird, P., "Minimizing conflicts: a heuristic repair method for constraint satisfaction and scheduling problems". *Artificial Intelligence*, 58, 161-205, 1992.

[14] Murphey, R.A., Pardalos, P.M., and Resende, M.G.C., "Frequency Assignment Problems", *AT&T Labs Research Technical Report*: 98.16.1, (to appear in Handbook of Combinatorial Optimisation, Kluwer Academic Publishers, 1999).

[15] Reeves, C.R., "Genetic Algorithms", in: Reeves, C. (ed.), *Modern Heuristic Techniques for Combinatorial Problems*, Blackwell Scientific Publishing, 151-196, 1993.

[16] vom Scheid, G., "Extension of GENET to Partial Constraint Satisfaction Problems and Constraint Satisfaction Optimisation Problems", *Master's Thesis*, Dept. of Mathematics, King's College London, UK, 1995.

[17] Selman, B., Levesque, H. J., and Mitchell, D.G., "A new method for solving hard satisfiability problems", *Proceedings of AAAI-92*, 440-446, 1992.

[18] Tiourine, S., Hurkins, C., and Lenstra, J.K., "An overview of algorithmic approaches to frequency assignment problems", *EUCLID CALMA Project Overview Report*, Delft University of Technology, The Netherlands, 1995.

[19] Tsang, E, *Foundations of Constraint Satisfaction*. Academic Press, 1993.

[20] Tsang, E., and Voudouris, C., "Fast local search and guided local search and their application to British Telecom's workforce scheduling problem", *Operations Research Letters*, Vol. 20, No. 3, 119-127, 1997.

[21] Voudouris, C., *Guided Local Search for Combinatorial Optimization Problems, PhD Thesis*, Department of Computer Science, University of Essex, Colchester, UK, July, 1997.

[22] Voudouris, C., and Tsang, E., "Partial Constraint Satisfaction Problems and Guided Local Search", *Proceedings of 2nd Int. Conf. on Practical Application of Constraint Technology* (PACT'96), London, April, 337-356, 1996.

[23] Voudouris, C., and Tsang, E., "Guided local search and its application to the travelling salesman problem", *European Journal of Operational Research*, Vol. 113, Issue 2, pp80-110, 1998.

[24] Voudouris, C., and Tsang, E., "Guided Local Search for the Radio Link Frequency Assignment Problem", *manuscript* to be submitted to the Journal of Heuristics.

[25] Wang, C.J., and Tsang, E., "Solving constraint satisfaction problems using neural-networks", *Proceedings of IEE Second International Conference on Artificial Neural Networks*, 295-299, 1991.

CELAR Set	Soluble Instances							Insoluble Instances					
Method	**1**	**2**	**3**	**4**	**5**	**11**	**Time**	**6**	**7**	**8**	**9**	**10**	**Time**
Simulated Annealing (EUT)	2	0	2	0	0	2	1min	6%	65%	5%	0%	0%	310min
Taboo Search (EUT)	2	0	2	0	-	0	5min	-	-	-	-	-	-
Variable Depth Search (EUT)	2	0	2	0	-	10	6min	3%	0%	14%	0%	0%	85min
Simulated annealing (CERT)	4	0	0	0	-	10	41min	42%	1299%	70%	2%	0%	42min
Tabu Search (KCL)	2	0	0	0	0	2	40min	167%	1804%	566%	8%	1%	111min
Extended GENET (KCL)	0	0	0	0	0	2	2min	12%	27%	40%	-	-	20min
Genetic Algorithms (UEA)	6	0	2	0	-	10	24min	0%	386%	134%	3%	0%	120min
Genetic Algorithms (LU)	-	-	-	-	-	-	-	0%	0%	0%	0%	0%	hours
GLS-FLS-S4	**0**	**0**	**0**	**0**	**0**	**2**	**2.91min**	**-1.34%**	**0.03%**	**0%**	**0%**	**0%**	**12min**
CALMA Project Best Solution	16*	14*	14*	46*	792*	22*		3437	343594	262	15571	31516	

Table 3 Comparison of the GLS with the heuristic search algorithms of the CALMA project on the Celar Instances.

GRAPH Set	Soluble Instances									Insoluble Instances						
Method	**1**	**2**	**3**	**4**	**8**	**9**	**10**	**14**	**Time**	**5**	**6**	**7**	**11**	**12**	**13**	**Time**
Simulated Annealing (EUT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Taboo Search (EUT)	0	0	0	0	0	0	398	0	3min	-	-	-	-	-	-	-
Variable Depth Search (EUT)	-	-	-	-	-	-	-	-	-	1%	0%	0%	0%	0%	0%	63min
Simulated annealing (CERT)	4	4	-	-	-	-	-	-	22min	-	123%	23%	-	-	-	81min
Tabu Search (KCL)	0	2	-	-	4	4	-	2	28min	409%	18%	303%	2648%	186%	1727%	minutes
Extended GENET (KCL)	0	0	-	-	2	4	-	-	5min	-	-	-	-	-	-	-
Genetic Algorithms (UEA)	2	2	i	i	-	10	i	4	9min	33%	282%	39%	763%	29%	342%	20min
Genetic Algorithms (LU)	-	-	-	-	-	-	-	-	-	0%	3%	0%	-	0%	104%	hours
GLS-FLS-S4	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**-2**	**1.29min**	**0%**	**-1.17%**	**0.14%**	**-12.12%**	**0%**	**-6.87%**	**5.46min**
CALMA Project Best Solution	18*	14*	380*	394*	20	18*	394	10		221	4189	4324	3513	11827	11130	

Table 4 Comparison of the GLS with the heuristic search algorithms of the CALMA project on the Graph Instances.

soluble instances
4 Number of frequencies above optimal solution
i infeasible solution reported

insoluble instances
42% deviation from the best known solution

- method not applicable or no solution is reported
* optimality of solution is proved
+ preprocessing time not included

Research Centers:

CERT - Centre d'Etudes et de Recherces de Toulouse, France;
DUT - Delft University of Technology, The Netherlands;
EUT - Eidhoven University of Technology, The Netherlands;
KCL - King's College London, United Kingdom;
LU - Limburg University, Maastricht, The Netherlands;
UEA - University of East Anglia, Norwich, United Kingdom.

Machines Used:
CERT - SUN SPARC 10
DUT - HP9000/720
EUT - SUN SPARC 4
KCL – DEC Alpha 3000 (130 Mhz)
LU (GA) - DEC OSF/1 AXP
LU(CS) -PC
UEA – DEC Alpha (133 Mhz)
GLS – UltraSparc (167 Mhz)

PAPER TITLE : **Paper 14 - Solving the Radio Link Frequency Assignment Problem Using Guided Local Search**

AUTHOR : C. Voudouris, E. Tsang

NAME : K.S. Kho

QUESTION :

In your results in Table 1 and 2, you reported solution cost varying from 14 to 393592. What were the scenarios used e.g. how many frequencies were to be assigned and what were the Frequency Assignment constraints?

ANSWER :

The problems were made available by the French Centre d'Electronique l'Armament, in a Web site maintained by Thomas Scliex before he moved on (http.//www-bia.inra.fr/T/Schiex/Doc/CELARE.html). The optimal values are not comparable between problems. The 25 benchmark problems cover problems of different classes, from satisfiable problems (in which cases the goal is to minimise the number of frequencies used or the maximum frequency used, e.g.) to insoluble problems (in which case other criteria are to be optimised). This is meant to cover a range of problems of the interest in the CALMA project.

I should perhaps point out that the GLS algorithm that we use is unchanged for all the test problems. In fact, the same GLS was used in other problems reported, with domain knowledge added and λ adjusted, of course.

PAPER TITLE : **Paper 14 - Solving the Radio Link Frequency Assignment Problem Using Guided Local Search.**

AUTHOR : C. Voudouris, E. Tsang

NAME : J. Bater

QUESTION :

My particular experience is with cellular problems. Given that local conditions, due to rapid fall off of signal power with distance, seem to predominate the structure of these problems. Do you think that the techniques you have described, for fast and guided band search which seem more able to capture these structures represent, the beginning of methods which combine good search algorithms with good modelling?

ANSWER :

This is a difficult question to answer, because we do not know when GLS-FLS is good and bad. This is something that we are trying to address. The "landscape" of the search space described by the questioner is a "hilly" one. (The modelling defines the landscape). We know that hill climbing in general is troubled by plateaux. So a hilly landscape favours GLS. Given that constraint violation usually occur locally though it can be propagated to other locations (i.e. other constraints), FLS would help the search to focus on the relevant parts of the landscape.

To summarise, the cellular problem has characteristics which favour GLS-FLS, though whether GLS-FLS is a start of a new class of algorithms for this application remains to be seen (depending op adoption by users and other researchers).

SOLVING THE RADIO LINK FREQUENCY ASSIGNMENT PROBLEM WITH THE GUIDED GENETIC ALGORITHM

T. L. LAU, and E. P. K. TSANG

University of Essex
Dept. of Computer Science
Wivenhoe Park
Colchester CO4 3SQ
United Kingdom
email: {tllau, edward}@essex.ac.uk

The Radio Link Frequency Assignment Problem is an abstraction of a real life military application that involves the assigning of frequencies to radio links. This problem set consists of eleven instances that are classed as either a Constraint Satisfaction Optimization Problem or a Partial Constraint Satisfaction Problem. Each problem has different optimization and constraint requirements, and can have up to 916 variables, and up to 5548 constraints.

The Guided Genetic Algorithm (GGA) is a hybrid of Genetic Algorithm and meta-heuristic search algorithm Guided Local Search. As the search progresses, GGA modifies both the fitness function and fitness template of candidate solutions based on feedback from constraints. In this paper, we have shown that GGA has the best optimality-robustness advantage over current published results.

Keywords: Genetic Algorithm, Constraint Satisfaction Optimization Problem, Partial Constraint Satisfaction Problem

1 Introduction

A finite Constraint Satisfaction Problem (CSP) can be described as a problem with a finite set of variables, where each variable is associated with a finite domain. Relationships between variables constraint the possible instantiations they can take at the same time [1, 2]. To solve a CSP, one must find the solution tuple that instantiate variables with values of their respective domains, and that these instantiations do not violate any of the constraints. Our area of research is in Constraint Satisfaction Optimization Problem (CSOP) and Partial Constraint Satisfaction Problem (PCSP), two variations of the CSP.

Paper presented at the RTO IST Symposium on "Frequency Assignment, Sharing and Conservation in Systems (Aerospace)", held in Aalborg, Denmark, 5-7 October 1998, and published in RTO MP-13.

In the realms of CSP, the instantiation of a variable with a value from its domain is called a *label*. A simultaneous instantiation of a set of variables is called a *compound label*, which is a set of labels. A *complete compound label* is one that assigns values, from the respective domains, to all the variables in the CSP.

A CSOP is a CSP with an objective function f that maps every complete compound label to a numerical value. The goal is to find a complete compound label S such that $f(S)$ gives an optimal value, and that no constraint is violated. A PCSP is similar to a CSOP except that the complete compound label may have variable instantiations that violate some of its constraints. Violation is unavoidable because the constraints are so tight that a satisfiable solution does not exist, or cannot be found [3, 1]. Deciding which constraint to violate is influenced by its cost and type. Hard constraints are types of constraints that must not be violated, whereas soft constraints may. The sum cost of all violated constraints is reflected in the objective function, further to other optimization criteria.

1.1 The Radio Link Frequency Assignment Problem

1.1.1 Background

The EUCLID CALMA (Combinatorial Algorithms for Military Applications) consortium is a group of six research bodies in Europe that was formed to investigate the use of AI techniques to aid military decisions. The Radio Link Frequency Assignment Problem (RLFAP) is a case study proposed within the group to observe the effectiveness of different approaches. It is an abstraction of a real world military application that is concerned with assigning frequencies to radio links. The RLFAP[1] contains eleven instances with various optimization criteria, made publicly available through the efforts of the French Centre d'Electronique l'Armament. RLFAP is NP-hard and is a variation on the T-graph colouring problem introduced in [4].

1.1.2 Types of Instances

Each instance in the RLFAP has a set of files that describe its variables, their domains, the constraints, and the objective. In addition, we are also given information on the respective optimization requirements based on the solubility of the problem. Optimization criteria describe the interpretation of variable instantiations and the means of measuring their desirability; thus shaping the objective function of our search routine, whereas the solubility of the problem states if a solution can be found under the condition that

[1]RLFAP is available at the Centre d'Electronique l'Armament (France), via ftp at ftp.cert.fr/pub/bourret.

no constraint was violated. If an instance can be solved without constraint violation, then its optimality is defined as either O1 or O2, otherwise it is O3 (see below). For an insoluble instance, we use the violation cost of each constraint to solve the instance as a PCSP. In this paper, regardless of the problem's solubility, all instances in RLFAP are solved as PCSP since a CSOP problem can be mapped into a PCSP by giving each constraint a violation cost. This violation cost is the same value for all constraints in the instance.

O1 - optimal solution is one with the fewest number of different values in its variables.

O2 - optimal solution is one where the largest assigned value is minimal.

O3 - if a the problem cannot be solved without violating constraints, find a solution that minimizes the objective function as follows:

$$\begin{aligned} &a_1 \times nc_1 + a_2 \times nc_2 + a_3 \times nc_3 + a_4 \times nc_4 + \\ &+ b_1 \times nv_1 + b_2 \times nv_2 + b_3 \times nv_3 + b_4 \times nv_4 \end{aligned} \tag{1}$$

Where nc_i is the number of violated constraints of priority i, nv_i is the number of modified variables with mobility i. Mobility for a radio link states the cost for changing the frequency from its assigned default. The values of the weights a_i and b_i are given if necessary.

All constraints in the RLFAP are binary; that is, each constraint operates on the values in two variables. These constraints test the absolute difference of two variables in a candidate solution, where this logical test can belong to either of the two following classes:

C1 - the absolute difference must be lesser than a constant.

C2 - the absolute difference must be equal to a constant.

Table 1 lists the instances, their characteristics and its objective. From this table, we can observe that the RLFAP contains instances that are varied in both the optimization and constraint criteria. Further, the number of variables, their domain sizes, and the number of constraints on these variables make the RLFAP a non-trivial problem set for any algorithm. The RLFAP would not only test the quality and robustness of an algorithm, but also its flexibility to adapt to the different optimization and constraint criteria of each instance.

Table 1: Characteristics of RLFAP instances.

Instance	No. of variables	No. of constraints	Souble	Minimize	Type
scen01	916	5548	Yes	Number of different values used	O1
scen02	200	1235	Yes	Number of different values used	O1
scen03	400	2760	Yes	Number of different values used	O1
scen04	680	3968	Yes	Number of different values used	O1
scen05	400	2598	Yes	Number of different values used	O1
scen06	200	1322	No	The maximum value used	O2
scen07	400	2865	No	Weighted constraint violations	O3
scen08	916	2744	No	Weighted constraint violations	O3
scen09	680	4103	No	Weighted constraint violations and mobility costs	O3
scen10	680	4103	No	Weighted constraint violations and mobility costs	O3
scen11	680	4103	Yes	Number of different values used	O1

2 The RLFAP in PCSP Expression

A PCSP is defined as a quadruple of $\{Z, D, C, f\}$ where Z is a finite set of variables. With respect to Z, D is a function that maps every variable to a set of values, which is called a domain. C is a finite set of constraints that affect a subset of the variables, and each constraint has a cost for its violation. The objective function f returns a magnitude based on the instantiation of the variables and the satisfaction of constraints. In the RLFAP, each instance has a set of files that conveniently describe the respective Z, D and C sets.

2.1 Variables and Domains

For any of the RLFAP instance with m variables, let q_j be a variable in Z representing one radio link. For each variable q_j in Z, there is one associated domain mapped by the function D, denoted by $D(q_j)$, which contains a set of n values, each value representing a valid frequency that can be assigned to the variable.

$$Z = \{q_1, q_2, \ldots, q_m\} \tag{2}$$

$$where\ \forall q_j \in Z: \quad D(q_j) = \{freq_1, freq_2, \ldots, freq_n\} \tag{3}$$

2.2 Constraints

The constraint set C consists of n elements, representing n constraints in the instance. Each element in C consist of the constraint c_i and its cost $cost_i$[2] (Eq. 4). As discussed in section 1.1.2, there are two types of binary constraints in the RLFAP; C1 and C2 which we formulate into Eq. 5. In that equation, q_a and q_b are two variables from a candidate solution, and z is a constant. Eq. 6 states that constraint c_i returns a binary value that is 1 for a violation and 0 otherwise.

$$C = \{< c_1, cost_1 >, < c_2, cost_2 >, \ldots, < c_n, cost_n >\} \quad (4)$$

$$\forall c_i \in C : \begin{cases} c_i \equiv |q_a - q_b| < z, \text{if } c_i \text{ is type C1} \\ c_i \equiv |q_a - q_b| = z, \text{if } c_i \text{ is type C2} \end{cases} \quad (5)$$

$$where\ c_i = \begin{cases} 1, \text{constraint is violated} \\ 0, \text{otherwise} \end{cases} \quad (6)$$

2.3 Objective Function

The respective objective function f of the instances in RLFAP are stated in Table 1. The objective functions are also explained in section 1.1.2.

3 Algorithms

CSPs and CSOPs are generally NP-hard [1] and although heuristics have been found useful in solving them, most systematic search algorithms are deterministic and constructive [5], and would thereby be limited by the combinatorial explosion problem. Systematic methods include search and inference techniques. These search methods are complete, so they are able to guarantee a solution, or to prove that one does not exist. Thus systematic techniques will, if necessary, search the entire problem space for the solution [6].

The combinatorial explosion is an obstacle faced by systematic search methods for solving realistic CSPs, and in looking for optimal and or near-optimal solutions in CSOPs. In optimization, to ensure that the solution found is the optimal, systematic search algorithms would need to exhaust the entire problem space to establish that fact.

Stochastic search methods are normally incomplete. They are not able to guarantee that a solution can be found, and neither can they prove that a solution does not exist. They forgo completeness for efficiency. Often, stochastic search methods can be faster in solving CSOPs than systematic

[2] Cost for violating the constraint.

methods [7]. Many publications such as [8, 9, 10] demonstrated on several large problems that systematic search algorithms fail to solve, but stochastic alternatives efficiently conquer.

3.1 Genetic Algorithms

Geneitc Algorithms are stochastic search algorithms that borrow some concepts from nature [11, 12, 13]. GA maintains a *population pool* γ of candidate solutions called *strings* or *chromosomes.* Each chromosome γ_p is a collection of α building blocks known as *genes*, which are instantiated with values from a finite domain. Let $\gamma_{p,q}$ denote the value of gene q in chromosome p in the population γ.

Associated with each chromosome is a *fitness* value which is determined by a user defined function. The function returns a magnitude that is proportional to the candidate solution's suitably and/or optimality. Fig. 1 shows the control and data flow of a canonical GA. At the start of the algorithm, an initial population is generated. Initial members of the population may be randomly generated, or generated according to some rules. The *reproduction operator* selects chromosomes from the population to be parents for a new chromosome and enters them into the *mating pool.* Selection of a chromosome for parenthood can range from a totally random process to one that is baised by the chromosome's fitness.

The *cross-over operator* oversees the mating process of two chromosomes. Two parent chromosomes are selected from the mating pool randomly and the *cross-over rate*, which is a real number between zero and one, determines the probability of producing a new chromosome from the parents. If the mating was performed, a child chromosome is created which inherits complementing genetic material from its parents. The cross-over operator decides what genetic material from each parent is passed onto the child chromosome. The new chromosome produced is entered into the *offspring pool.*

The *mutation operator* takes each chromosome in the offsrping pool and randomly change part of its genetic make-up, ie. it's content. The probability of mutation occuring on any chromosome is determined by the user specified mutation rate. Chromosomes, mutated or otherwise, are put back into the offspring pool after the mutation process.

Thus each new generation of chromosomes are formed by the action of genetic operators (reproduction, cross-over and mutation) on the older population. Finally, the members of the population pool are compared with those of the offspring pool. The chromosomes are compared via their fitness value to derive a new population, where the weaker chromosomes may be eliminated. In exact, weaker members in the population pool is replaced by the fitter child chromosomes from the offspring pool. The heuristic for

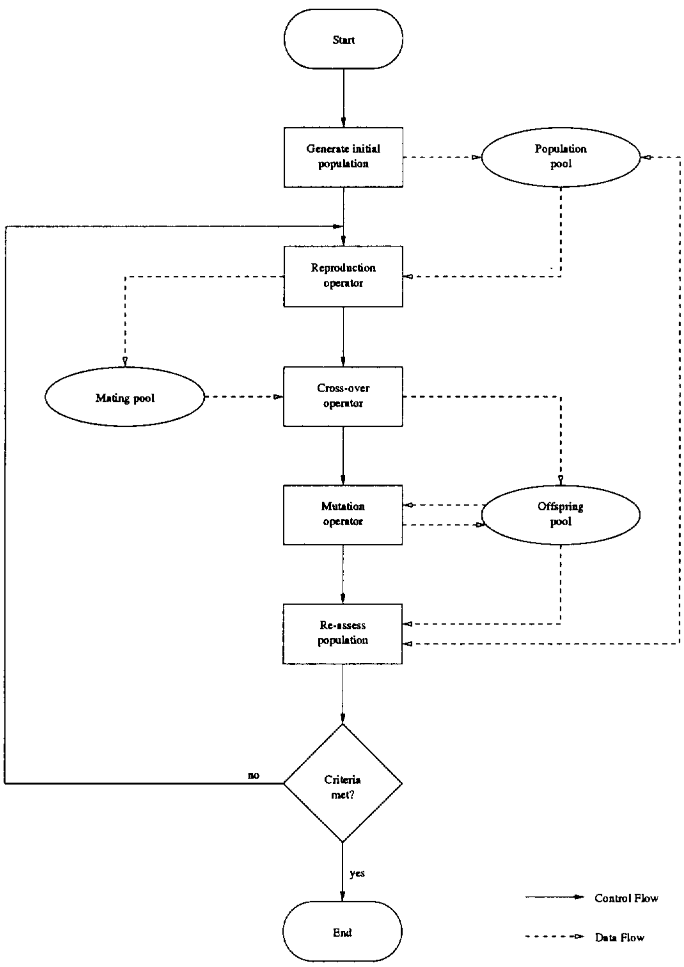


Figure 1: A canonical Genetic Algorithm

assessing the survival of each chromosome into the next generation is called the *replacement strategy*.

The process of reproduction, cross-over, mutation and formation of a new poplution completes one *generation* cycle. A GA is left to progress through generations, until certain criteria (such as a fixed number of generations, or a time limit) are met. GAs were initially used for machine learning systems, but it was soon realised that GAs have great potential in function optimization [14, 12, 11].

3.1.1 Shortcomings of GAs when solving CSPs

In applying the canonical GA to solve instances belonging to the CSP class, the problem of *high epistasis* often limits its success. Epistasis is the interaction between different genes in a chromosome. A candidate solution to a typical CSOP is often represented as a chromosome, where each gene in the chromosome describes a variable in the CSOP. Constraints influence both the values that sets of genes can take simulanteously and the overall fitness of that chromosome. Goldberg suggested high epistasis as an explanation to GAs failure in certain tasks [12].

3.2 Guided Genetic Algorithm

3.2.1 Background

Among our eariler work on CSOPs, we looked at the Processor Configuration Problem (PCP) [15, 16]. Briefly, the PCP is a real life CSOP where the task is to link up a finite set of processors into a network, whilst minimizing the maximum distance between these processors. Since each processor has a limited number of communication channels, a carefully planned layout will help reduce the overhead for message switching.

We developed a GA called the Lower Power Genetic Algorithm (LPGA) [17, 18] specifically for solving the PCP. LPGA is a two-phase GA approach where in the first phase, we run LPGA until a local optimum has been determined. The best chromosome from this run is analysed and used to construct a fitness template for use in the next phase. This fitness template is a map that defines undesirable genes, so influencing LPGA to change their contents. By insisting that crucial genes do not change, the evolution in the second phase shifts focus onto other parts of the string; resulting in a more compact search space.

LPGA found solutions better than results published so far in the PCP. It's success could be attributed to the use of an effective data representation and more importantly, the presence of an application specific penalty algorithm. In our effort to generalize LPGA, we sought to develop a GA that utilizes a dynamic fitness template constructed by a general penalty algorithm. The Guided Genetic Algorithm (GGA) repeated in this paper was the result of this effort.

3.2.2 Overview of GGA

In our jounrey to develop GGA, we have taken liberty with some of the traditional GA concepts (such as the addition of a penalty operator, and an alternate interpretation of the mutation rate). These will be introduced as we progress through the rest of this paper. Comparing GGA in Fig. 2 against the canonical GA in Fig. 1, we could see the additions of a data collection called the fitness templates, a penatly operator (see 3.2.3) and a condition to activate that operator. Also added to Fig. 2 are the interactions between the penalty operator and the data space in GGA. Appended at the end of this section (section 3) are two tables (Table 2 and 3), summarizing for the readers' convenience, the terms and technology introduced henceforth.

The control flow of the GGA is very much similar to that of the canonical GA, described in section 3.1. After the start of the algorithm, an initial population is created. A new generation of chromosomes are derived from the parent chromosomes through the actions of the reproduction, cross-over and mutation operators. Both the cross-over and mutation operators (or any operator thereof) may be adapted to use the information provided by the penalty operator, via the *fitness template* of each chromosome which is collected in the *fitness templates* (explained in section 3.2.4). Memberships to the population pool are re-assessed by comparing the fitness of the chromosomes from the population and offspring pool. For the RLFAP, GGA was configured to use an *elitist replacement strategy.* Under this strategy, chromosomes from both the population and offspring pool are ranked by their fitness. The fittest n chromosomes in the ranking are used to form the next generation's population pool. In GGA, n is set to the size of the population pool.

New elements of the GGA comes into play at this point. The new population is surveyed for the possibility of being trapped in a local optimum. We can observe that when a search is trapped in a local optimum, it repeatedly returns the same solution since the neighbouring states does not offer any improvement. If the population is indeed trapped in a local optimum, the penalty operator is called. The penalty operator looks for undesirable features in the chromosomes and update the fitness template (or fitness templates), so that mutation and cross-over operators might fade out these features in the coming generations.

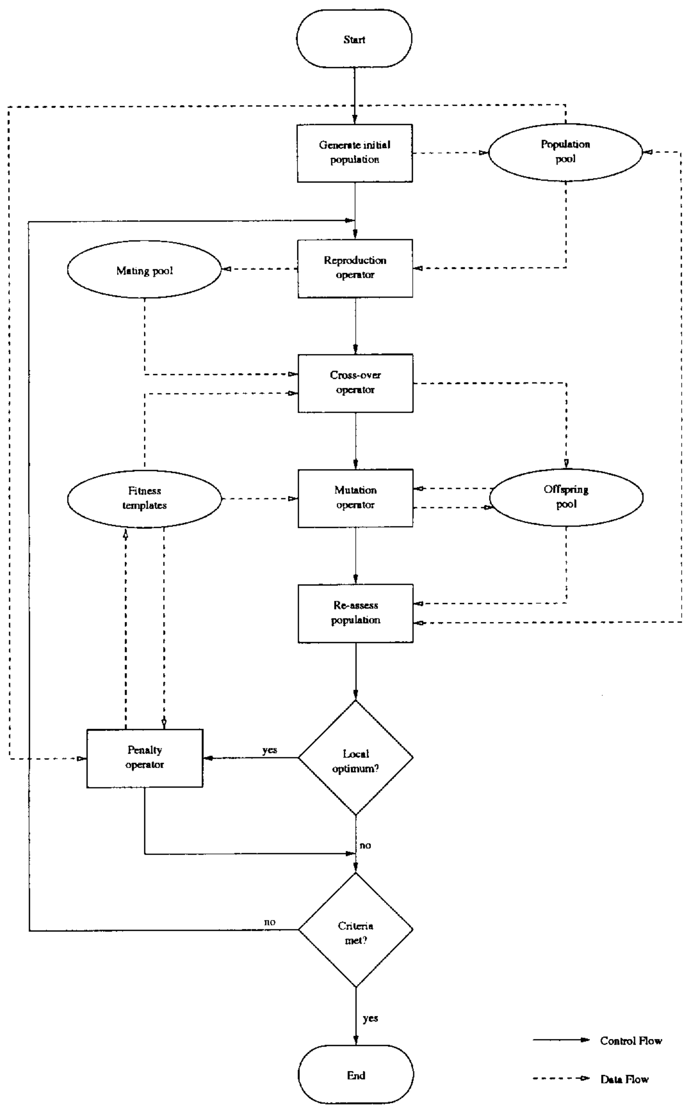


Figure 2: The Guided Genetic Algorithm

3.2.3 Penalty Operator

In LPGA, its use of a fitness template (generated by a specialized penalty algorithm) was the motivating force in the development of GGA. In the quest for a general penalty algorithm, we looked for functional similarities to LPGA and more importantly, that the nature of the penalty algorithm will not be obstructive to the operation of a canonical GA. The Guided Local Search (GLS) [19] developed by our research group is an intelligent search scheme for combinatorial optimization problems. It met our criteria and further, its conceptual simplicity and proven effectiveness in a range of well known problems was an added attraction [20, 21, 22, 23]. In GGA, we adapted GLS in the form of the *penalty operator*.

Solutions are characterized by a set of solution features θ, where a solution feature θ_i can be any property exhibited by the solution (Eq. 7). This property must be non-trival, such that it does not appear in all candidate solutions. Research on GLS has indicated that feature definition is not difficult, since the domain often suggests features that one could use. The application in this paper supports this point.

$$\theta = \{\theta_1, \theta_2, \ldots, \theta_m\} \tag{7}$$

In GGA, a fearure is limited to variable assignments (in GLS, it is more general); a feature in a chromosome may be exhibited by the simultaneous assignments of a group of genes. Thus the feature θ_i defines a set of positions in the chromosome representation. And the feature θ_i is represented by an *indicator function* τ_i in Eq. 8, which test the existence of that feature. For each feature θ_i, there is a *cost* η_i which rates that feature's presence in a solution in degrees of undesirability. Indicator functions and costs are application dependent, and so they are defined by the user.

$$\tau_i(\gamma_p) = \begin{cases} 1, & \text{solution } \gamma_p \text{ exhibits feature } \theta_i \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

Penalty counter ζ_i is a variable maintained by GGA that gives the degree of extent that the feature θ_i is penalized as the search is progressing; the counter is initialized to zero at the beginning of the search. A new fitness function called the *augmented cost function* g (Eq. 9) is used in place of function f, so that changes in penalty counters will affect the survival of chromosomes. The *regularization parameter* λ (adopted from GLS) measures the impact penalties have, with respect to function f.

$$g(\gamma_p) = f(\gamma_p) + \lambda \cdot \sum(\zeta_i \cdot \tau_i(\gamma_p)) \tag{9}$$

In GGA, if the fitness of the best chromosome remains unchanged for a specific number of generations, we conclude that it is trapped in a local optimum. The penalty operator comes in to analyze the best chromosome γ_p of the population for features to penalize. Penalties are used in GGA to guide the search to escape local optima. To evaluate the utility of penalizing individual features exhibited by a candidate solution, GGA (following GLS) takes into consideration the cost as well as the penalty counter (Eq. 10). Thus, for all features θ_i in the fittest chromosome γ_p that maximizes the function $util(\gamma_p, \theta_i)$ (Eq. 10), the related penalty counter ζ_i is incremented by one. It is hoped that by penalizing undesirable features, we can escape from the local optimum and suppress the occurance of these features in the coming generations.

$$util(\gamma_p, \theta_i) = \tau_i(s) \cdot \frac{\eta_i}{1 + \zeta_i} \tag{10}$$

3.2.4 Fitness Template

Central to the theme in GGA is the fitness templates. Besides the fitness function, the fitness templates offer an added channel of communication between the penalty operator, and the mutation and cross-over operators. The fitness template is a map that defines which genes in a chromosome are more susceptible to be changed during cross-over or mutation.

In GGA, each chromosome γ_p in the population is associated with exactly one *fitness template* δ_p. Each fitness template is made up of smaller units known as *weights* $\delta_{p,q}$, each of which corresponds to a gene $\gamma_{p,q}$. A weight $\delta_{p,q}$ is a positive integer. The "heavier" a gene appears (compared to its comrades), the greater are its chances of having its content altered. Therefore in the case of mutation, the weight of a gene is proportional to the probability that mutation may occur on it, relative to the weights of other genes in the same chromosome. This is especially useful when the number of genes in a chromosome is large, where random selection of genes might not be helpful. More details on the role of the fitness template with the mutation and cross-over operator will be given in their respective sections.

Weights in the fitness template for each chromosome are computed when the chromosome was first created, and after the penalty operator has penalized feature(s). Computation of weights are needed after these events because the content of either the chromosomes or the penalty counters have changed.

For a chromosome, the distribution process (Fig. 3) starts by initializing all weights to zero. It will check the chromosome for the presence of any features from the set θ. For a feature θ_i that exist in the chromosome[3], all

[3]Feature θ_i is present when its indicator function τ_i returns a one.

the weights related to the gene positions defined by θ_i is incremented with the value in its penalty counter ζ_i.

```
FUNCTION DistributePenalty( chromosome γ_p )
{
    FOR EACH weight δ_{p,q} RELATED TO chromosome γ_p
    {
        δ_{p,q} ← 0
    }

    FOR EACH solution feature θ_i IN feature set θ
    {
        IF τ_i(γ_p) = 1 THEN
        {
            FOR EACH gene position q defined by θ_i
            {
                δ_{p,q} ← δ_{p,q} + ζ_i
            }
        }
    }
}
```

Figure 3: Algorithm for the Distribution of Penalty

3.2.5 Cross-over Operator

The action of mating two individuals from the population produces a new child. Each parent contributes a set of genes which the child inherits. In GA, the process of choosing parents, deciding their respective contribution rights of genetic material, and the forging of a child chromosome from these material is the responsibility of the cross-over operator. The probabilty of cross-over occuring is controlled by the parameter *cross-over rate*. By assembling a new chromosome that contains parts of two parent chromosomes, it may introduce to the popultion a new point in the search space. And since the parents chosen for mating are selected with bias to their fitness, we hope that the child chromosome may be fitter.

Cross-over operators differ primarily from each other in the way that they choose the genes from the parents to form the child. In the canonical GA, one of the simplest form of cross-over is the *one-point cross-over* [12, 11]. In Fig. 4, we have two parent chromosomes whose genes are binary encoded. One random point along the length of the chromosomes are selected as the cross-over point. Each parent donates one different part of their chromosome (defined by the cross-over point) to create the child chromosome.

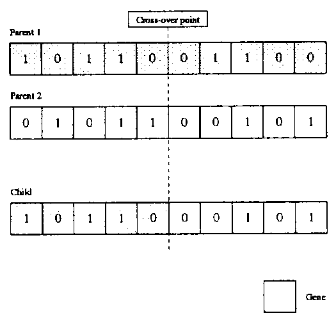


Figure 4: An example of the One-point Cross-over Operator in action

In GGA, we have adapted the cross-over operator to take advantage of the fitness template. Two chromosomes γ_p and $\gamma_{p'}$ are selected as parents to produce the child $\gamma_{p''}$. Each gene in chromosome γ_p competes against the corresponding gene in $\gamma_{p'}$ for a place in the child. This compeitition is a weighted random selection, influenced by the weights $\delta_{p,q}$ and $\delta_{p',q}$ of the respective genes; thus the "lighter" gene will have a greater chance to propagate its information to the child. Note that the child does not inherit the weights from its respective parent, since the child may represent a different solution from its parents, and thus requiring the penalty operator to re-assess it. The algorithm of GGA's cross-over operator is shown in Fig. 5, and Fig. 6 shows its effect when applied to the situation for one-point cross-over in Fig. 4.

The operator starts off by receiving two parents γ_p and $\gamma_{p'}$ from the mating pool. For each set of corresponding genes $\gamma_{p,q}$ and $\gamma_{p',q}$ in the parents, it computes the *sum* of their weights. The selection of the gene is randomly biased, such that the probability for either $\gamma_{p,q}$ or $\gamma_{p',q}$ to have its gene passed onto their child is $\frac{\gamma_{p',q}}{sum}$ and $\frac{\gamma_{p,q}}{sum}$ respectively; giving the advantage to a "lighter" gene, which we would want the child $\gamma_{p'',q}$ to inherit. This gene selection process is repeated for all genes in the parents. When a child chromosome is complete, its fitness and weights are computed.

3.2.6 Mutation Operator

Mutation produces variations in the population through altering the information that genes carry. The *mutation rate* states the probability that

FUNCTION **CrossOver**(parent chromosomes γ_p and $\gamma_{p'}$)

{

FOR EACH gene position q IN the chromosome

{

$sum \leftarrow \delta_{p,q} + \delta_{p',q}$

$point \leftarrow$ random integer from $\{0, \ldots, sum - 1\}$

IF $point < \delta_{p,q}$ THEN

{

$gene \leftarrow \gamma_{p',q}$

}

ELSE

{

$gene \leftarrow \gamma_{p,q}$

}

}

RETURN gene as $\gamma_{p'',q}$ for the offspring

}

Figure 5: Algorithm of the GGA Cross-over Operator

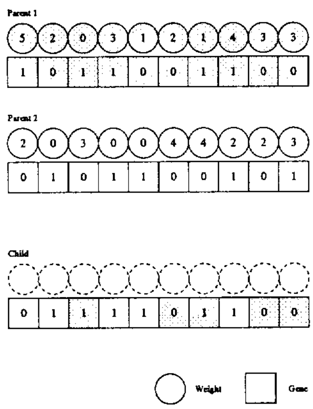


Figure 6: An example of the GGA Cross-over Operator in action
(Weights for the child is calculated afresh, not inherited)

mutation may occur on a chromosome. In GGA, the mutation rate is defined as a fraction of the size of each chromosome; the number of genes in a chromosome to mutate is the product of the mutation rate and the size of that chromosome.

In GGA, mutation (Fig. 7) acts on every child chromosome $\gamma_{p''}$ produced by the cross-over operator. For each chromosome, a number of genes are chosen (as above, decided by the mutation rate) to be modified. A gene $\gamma_{p'',q}$ is selected using the *roulette wheel selection method.* In this selection method, the probability for each gene to be picked is directly proportional to its weight. Thus a gene with a "heavier" weight (and therefore less desirable) compared to others in the chromosome, will have a greater chance of being selected. Appended below is a description of our implementation of the roulette wheel selection method.

> Given the chromosome $\gamma_{p''}$, we compute the sum of all the weights in the fitness template associated to this chromosome as $sum = \sum_{q=1}^{\alpha} \delta_{p'',q}$. The probabilty that a gene $\gamma_{p'',q}$ is selected is proportional to its weight over *sum*, ie. $P(\gamma_{p'',q}) = \frac{\delta_{p'',q}}{sum}$.

The next step for a selected gene $\gamma_{p'',q}$ is to seek an appropriate value for replacement. This could be totally random or in GGA's case, a value that will derive the best fitness (the biggest improvement) for the chromosome. In our algorithm, we have the variables *best* and *list*. The variable *best* holds the best fitness value, while *list* contains a list of values that will allow the chromosome to arrive at the fitness value in *best*. We step through all the values x_j in the domain $D(q)$ relevant to the gene $\gamma_{p'',q}$. If x_j produces a new fitness z greater[4] than *best*, *best* is set to z and *list* is emptied. However, if the z is equal to *best*, the value x_j is added to the *list*. When all values in the domain have been exhausted, we randomly instantiate $\gamma_{p'',q}$ with a value from *list*. Since at this point, *list* should contain all possible values that will give the chromosome $\gamma_{p''}$ the biggest improvement.

Updating of a gene's weight takes place after it's value has changed, where the weight associated with it is reduced by one unit [5] so that the probability of the same gene getting selected by the roulette wheel selector is reduced.

Gene mutation is repeated until the stopping criteria is met. As stated before, we stop mutating when the number of genes changed have reached a value that is the product of the mutation rate and the chromosome's length.

[4]Since RLFAP is a minimization problem, we would want the greatest decent.

[5]Since a weight is a positive integer, a weight will only be decremented if it is greater than zero.

```
FUNCTION Mutation( chromosome γ_p'' )
{
    i ← 0

    WHILE i < mutation rate × α (length of chromosome)
    {
        q ← RouletteWheel(γ_p'')
        best ← g(γ_p'')
        list ← γ_p'',q

        FOR EACH value x_j IN domain D(q)
        {
            γ_p'',q ← x_j
            z ← g(γ_p'')

            IF z ≥ best THEN
            {
                IF z > best THEN
                {
                    best ← z
                    list ← {}
                }

                list ← list + x_j
            }
        }

        γ_p'',q ← random value in list
        i ← i + 1
    }

    RETURN the mutated chromosome γ_p''
}
```

Figure 7: Algorithm of the GGA Mutation Operator

Table 2: Components of GGA

Algorithms	Purpose
Cross-over operator	Uses the fitness templates of two parent chromosomes to decide each parent's contribution of genetic material towards creating a child chromosome.
Mutation operator	The fitness template of a chromosome is used to guide in the alteration of the chromosome's genetic content.
Penalty operator	This operator detects and selects undesirable solution features in a chromosome to penalize. Penalization involves incrementing penalty counters of the assocaited features.
Local optimum detector	Detects if the search is trapped in a local optimum. If it is, the penalty operator is called.
Distribute Penalty	If a solution feature is present in a chromosome, the penalty counter associated with this feature is added onto the weights of the genes that are constituents of this feature.

Data structures	Purpose
Weight δ	Each gene has one weight. The weight is a measure of undesirability of the gene's current instantiation, compared to the rest of the chromosome.
Fitness template	A fitness template is a collection of weights. Each chromosome has one fitness template.
Solution feature θ	Solution features are domain specific and user defined. A feature is exhibited by a set of variable assignments that describes a non-trival property of a problem.
Penalty counter ζ	Each feature has one penalty counter. A penalty counter keeps count of the number of times its related solution feature has been penalized since the start of the search.

Table 3: Inputs and Parameters to GGA

Inputs/Parameters	Purpose
Solution feature θ	See Table 2
Cost η	Each solution feature has a cost to rate its undesirability of presence.
Indicator function τ	Each solution feature has a user defined indicator function that tests for the feature's presence in a chromosome.
Objective function	A function that maps each solution to a numerical value.
Regularization parameter λ	A parameter that determines the proportion of contribution that penalties have in an augmented fitness function.
Augmented fitness function g	A function that is the sum of the objective function on a chromosome and the penalties of features that exist in it.
Mutation rate	A fraction that defines the number of genes in the chromosome to mutate.
Cross-over rate	The probability that cross-over will occur between two chromosomes.

4 Preparing GGA to solve RLFAP

In section 2, we expressed the RLFAP as a formal PCSP. In this section, we discuss the steps needed to adapt those definitions into a form that GGA can use.

The feature set θ is a union of the feature set of constraints θ_{cst} and the set of mobility of radio links θ_{mbt} (Eq. 11). Constraint c_i defined in Eq. 5 is recast as a feature in the set θ_{cst} (Eq. 12), where a one is returned if the constraint cannot be satisfied, and zero otherwise. The value of cost ζ_{cost_i} to each constraint c_i depends on the nature of the instance. If the instance is soluble, then all ζ_{cst_i} are set to a large value; usually 10000, to signify that the constraint must not be borken (ie. hard constraints). For insoluble instances, ζ_{cst_i} is set to the weights given for its priority class (see section 1.1.2). Similar to soluble instances, hard constraints in insoluble instances will have their ζ_{cst_i} set to a large value. The set θ_{cst} has n features, where n is the number of constraints in the instance.

For the O3 objective type of instances, we need to minimize the mobility cost of our candidate solution, in addition to minimizing constraints violation costs. The set of mobility cost defines our next feature set, θ_{mbt} (Eq. 13). For each variable in these instances, there is a mobility cost ζ_{mbt_i} and a default assigned frequency $default_i$. If in our candidate solution, a variable has been assigned a value different from its default $default_i$, then a one is returned and zero otherwise. The mobility cost ζ_{mbt_i} is set to the weights given for its prioriy class (again see section 1.1.2). There are radio links whose frequency should never change, and the mobility cost for these have been set to a large value. The feature set θ_{mbt} has n features, where n is the number of variables in the instance.

$$\theta = \{\theta_{cst}, \theta_{mbt}\} \tag{11}$$

$$\forall \theta_i \in \theta_{cst} : \tau_{cst_i}(\gamma_p) \equiv \begin{cases} 1, \text{if C1 and } |\gamma_{p,a} - \gamma_{p,b}| \geq z_i \\ 1, \text{if C2 and } |\gamma_{p,a} - \gamma_{p,b}| \neq z_i \\ 0, \text{otherwise} \end{cases} \tag{12}$$

$$\forall \theta_i \in \theta_{mbt} : \tau_{mbt_i}(\gamma_p) \equiv \begin{cases} 1, & \text{if } \gamma_{p,i} \neq default_i \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

For all instances in the RLFAP, we seek to minimize the function g (Eq. 14). In g, the function f depends on the objective type of the instance (Eq. 15). The cost ζ_i and τ_i both refers to the unified feature set of θ. They will automatically associate with ζ_{cst_i} and τ_{cst_i}, or ζ_{mbt_i} and τ_{mbt_i} where applicable. Thus the value of n in Eq. 14 is the sum of number of features in θ_{cst} and θ_{mbt}.

$$g(\gamma_p) = f(\gamma_p) + \lambda \cdot \sum_{i=1}^{n}(\zeta_i \cdot \tau_i(\gamma_p)) \quad (14)$$

$$f(\gamma_p) = \begin{cases} \text{if O1,} & \text{number of different values used in } \gamma_p \\ \text{if O2,} & \text{largest value used in } \gamma_p \\ \text{if O3,} & a_1 \times nc_1 + a_2 \times nc_2 + a_3 \times nc_3 + a_4 \times nc_4 + \\ & + b_1 \times nv_1 + b_2 \times nv_2 + b_3 \times nv_3 + b_4 \times nv_4 \end{cases} \quad (15)$$

5 Benchmark

The RLFAP benchmark results by algorithms devised within the CALMA group was reported by Tiourine et al. in [24]. In this section, we compare GGA's results with the CALMA algorithms (section 5.2). Further, we will also evaluate the examine the value that GGA adds to the canonical GLS (section 5.3).

5.1 Test Environment

In our physical environment, GGA was written in C++ and complied using GNU GCC version 2.7.1.2. The code runs on an IBM PC compatible with a Pentium 133 MHz processor, 32MB of RAM and 512KB of Level 2 cache. Both compilation and executaion of GGA was performed on the Linux operating system, using kernel version 2.0.27. Under GGA's environment, we have a mutation rate and cross-over rate of 1.0, a population size of 20, and a stopping criterion of 100 generations. For the fitness function g, λ has a value of 10.

5.2 Comparing Quality of Solutions

In Table 4, we see the published results of all the CALMA algorithms, GLS and GGA. Algorithms from the CALMA project groups consist of either complete or stochastic methods. The results recorded in the table are from the best solution each algorithm had generated. For soluble instances (scen01, scen02, scen03, scen04, scen05 and scen11)[6], we report the number of frequencies above the known optimum that each solution (generated by the respective algorithms) uses. Results for the insoluble instances are reported as the percentage deviation from the best known reported solution.

For soluble instances, we observe that only seven out of the 13 algorithms[7] were able to provide a solution to all the instances. Of the six

[6] Instance scen06 was found to be insoluble, and thus solved as an O3 problem.

[7] Genetic Algorithms (LU) was designed for PCSP, and so it did not attempt any of the soluble instances.

soluble instances, GGA failed to return an optimum solution only for instance *scen11*. This instance has proved to be difficult for most of the algorithms, since only three algorithms (Taboo Search (EUT), Branch and Cut (DUT,EUT) and Constraint Satisfaction (LU)) were able to return a solution on par with the best known. Only two algorithms (Branch and Cut (DUT,EUT) and Constraint Satisfaction (LU) were able to report solutions that gave the most optimal resutls. However, these two algorithms were limited to solving soluble instances only.

Looking at insoluble instances, we see that 11 out of the 14 algorithms were able to solve these PCSP instances. Of the 11, nine of these algorithms managed to provide a satisfactory solution to the instances. Of note is the Genetic Algorithms (LU), which has the best known solutions.

Overall, we see that top performers in each category (soluble and insoluble) are limited to only that category; Branch and Cut (DUT,EUT) and Constraint Satisfaction (LU) for soluble instances, and Genetic Algorithms (LU) for insoluble instances. Of the total of 14 algorithms, 10 were applicable to both categories. Of these 10, only four algorithms were able to provide a satisfactory solution to all the 11 instances. Of the four, GLS and GGA consistently gave solution quality very close to the best known solutions.

5.3 Comparing GGA and GLS

From Table 4, we see that solution quality reported by GLS and GGA were very much the same except in soluble instance *scen11*, where GGA managed to better GLS's result. The amount of CPU time required for computing these results have shown GLS to be much superior. However, it is unfair to view GGA as just a parallel version of GLS. In section 3.2, we describe the mechanics of GGA. We have integrated GLS as a component of GGA; ie. as the penalty operator. The feedback from the GLS is used at two levels within GGA. On one level, GLS modifies the objective (fitness) function of GGA to influnce its search. On another level, information from GLS are encoded into the template of each chromosome, which rates the relative fitness of each gene in the chromosome. The templates are used to influence the cross-over and mutation processes. But because GGA manages several candidate solutions at the same time, it will have a more complex (and perhaps less effective) means to detect local optimum traps.

Besides solution quality and computation speed, one other measure is robustness. Robustness of an algorithm measures the consistency of the solutions it returns. In certain environments, one may need to be absolutely sure that the solution returned by an algorithm is, or very near the optimum. Following, we compare the statistics of GGA and GLS.

For the comparison, GGA runs with a population of only five chromo-

Table 4: Comparison of GGA with GLS and the CALMA project algorithms.

	Soluble Instances							Insoluble Instances						
Instance (scen)	01	02	03	04	05	11	Time	06	07	08	09	10	Time	Platform
Simulated Annealing (EUT)	2	0	2	0	0	2	1min	6%	65%	5%	0%	0%	310min	SUN Sparc 4
Taboo Search (EUT)	2	0	2	0	-	0	5min	-	-	-	-	-	-	SUN Sparc 4
Variable Depth Search (EUT)	2	0	2	0	-	10	6min	3%	0%	14%	0%	0%	85min	SUN Sparc 4
Simulated Annealing (CERT)	4	0	0	0	-	10	41min	42%	1299%	70%	2%	0%	42min	SUN Sparc 10
Tabu Search (KCL)	2	0	0	0	0	2	40min	167%	1804%	566%	8%	1%	111min	DEC Alpha
Extended GENET (KCL)	0	0	0	0	0	2	2min	12%	27%	40%	-	-	20min	DEC Alpha
Genetic Algoritms (UEA)	6	0	2	0	-	10	24min	0%	386%	134%	3%	0%	120min	DEC Alpha
Genetic Algorithms (LU)	-	-	-	-	-	-	-	0%	0%	0%	0%	0%	hours	DEC Alpha
Partial Constraint Satisfaction (CERT)	4	0	6	0	0	-	28min	83%	2563%	246%	47%	12%	6min	SUN Sparc 10
Potential Reduction (DUT)	0	0	2	0	0	-	3min	27%	-	-	4%	1%	10min	HP 9000/720
Branch and Cut (DUT,EUT)	0	0	0	0	0	0	<10min	-	-	-	-	-	-	-
Constraint Satisfaction (LU)	0	0	2	0	0	0	hours	-	-	-	-	-	-	PC
Guided Local Search (UE)	0	0	0	0	0	6	20sec	4%	9%	7%	0.7%	0.003%	2.88min	DEC Alpha
Guided Genetic Algorithm (UE)	0	0	0	0	0	2	30min	4%	9%	7%	0.7%	0.003%	2.88min	PC Linux
Best known solution	16	14	14	46	792	22		3437	343594	262	15571	31516		

Results for the soluble instances are reported as the number of frequencies more than the optimum used.
Results for the insoluble instances are reported as the percentage deviation from the best known solution.

CERT Centre d'Etudes et de Recherces de Toulouse, France
DUT Delft University of Technology, The Netherlands
EUT Eidhoven Univeristy of Technology, The Netherlands
KCL King's College London, United Kingdom
LU Limburg University, Maastricht, The Netherlands
UEA University of East Anglia, Norwich, United Kingdom
UE University of Essex, United Kingdom

somes[8]. We introduce variation of GLS, called *GLS5* to compete with GGA on even grounds. GLS5 is five GLS running concurrently of each other, each maintaining its own candidate solution. Run time and iteration count for each GLS thread within GLS5 has also been extended to meet with GGA's. At the end of each run, only the best solution from GLS5 was used. For all 11 instances, the three algorithms (GGA, GLS and GLS5) were each sampled 50 solutions for each instance.

The results from Table 5 shows GGA to have better robustness than GLS or GLS5 in soluble instances. Whereas results for insoluble instances are mixed, with both algorithms having results very close to each other.

6 Conclusion

GGA was devised as a GA for arresting the effect of high epistatis when GAs are deployed to solve problems such as those from the CSP class. In the benchmarks, we have shown that GGA adds value to the canonical GLS. And that overall, GGA performed well against the other algorithms. As a GA, GGA was more flexible than Genetic Algorithms (LU), and performed better than Genetic Algorithms (UEA).

The integration of GLS and the introduction of new elements to the foundation of the canonical GA gave GGA a technique of measuring gene fitness for a chromosome, and the provision for multi-criteria optimization. By knowing a gene's fitness within a chromosome, one could understand the magnitude of its contribution to the overall fitness. Gene fitness influence the effects that genetic operators have on them, encouraging change to genes with low fitness, whilst protecting the healthy ones.

For most applications where the users are more concerned with turnaround time and less so on robustness, GLS is clearly the better choice. But for mission critical applications, or applications where time is not as tight, GGA can guarantee robustness and may perform better than GLS at times.

References

[1] E. P. K. Tsang, *Foundations of Constraints Satisfaction.* Academic Press Limited, 1993.

[2] E. Freuder and A. Mackworth, *Constraints-based Reasoning.* MIT Press, 1994.

[8]This is the minimum number of chromosome needed to maintain any advantage over GLS.

Table 5: Comparing Robustness between GGA, GLS and GLS5.

Instance	Best Solution	Average Cost			Standard Deviation			Worst Solution		
		GLS	GLS5	GGA	GLS	GLS5	GGA	GLS	GLS5	GGA
scen01	16	18.6	17.0	16.0	2.3	0.8	0.0	22	18	16
scen02	14	14.0	14.0	14.0	0.0	0.0	0.0	14	14	14
scen03	14	15.4	14.4	14.0	1.3	0.4	0.0	18	16	14
scen04	46	46.0	46.0	46.0	0.0	0.0	0.0	46	46	46
scen05	792	792.0	792.0	792.0	0.0	0.0	0.0	792	792	792
scen06	3628	4333.8	4029.6	4029.6	766.0	538.2	529.3	6042	6028	6028
scen07	427054	530641.1	510532.5	51344.8	79666.7	75149.4	75612.8	700685	694011	698837
scen08	294	335.7	322.6	320.5	34.7	23.1	21.5	377	372	368
scen09	15805	15999.7	15895.0	15889.6	194.7	112.6	109.3	16340	16280	16124
scen10	31533	31686.6	31621.4	31626.1	146.1	101.7	104.8	31942	31922	31922
scen11*	28	-	33.9	30.2	-	3.1	1.7	-	36	32

* For *scen11*, results for GLS could not be computed because it did not return a satisfiable solution for some runs.

[3] E. Freuder and R. Wallace, "Partial constraint satisfaction," *Artificial Intelligence*, vol. 58, pp. 21–70, 1992.

[4] K. Hale, "Frequency assignment: Theory and applications," in *IEEE*, vol. 68, pp. 1497–1514, 1980.

[5] P. Meseguer, "Constraint satisfaction problems: An overview," *AI Communications*, vol. 2, pp. 3–17, Mar. 1989.

[6] Kumar, "Algorithms for constraint satisfaction problems: A survey," *AI Magazine*, vol. 13, no. 1, pp. 32–44, 1992.

[7] E. Freuder, R. Dechter, B. Selman, M. Ginsberg, and E. Tsang, "Systematic versus stochastic constraint satifaction," in *14th International Joint Conference on Artificial Intelligence*, 1995.

[8] S. Minton, M. Johnston, A. Philips, and P. Laird, "Minimizing conflicts: A heuristic repair method for constraint satisfaction and scheduling problems," *Artificial Intelligence*, vol. 58, pp. 161–205, 1992.

[9] B. Selman, H. Levesque, and D. Mitchell, "A new method for solving hard satisfiability problems," in *10th National Conference on Artificial Intelligence*, pp. 440–446, 1992.

[10] P. Langley, "Systematic and non-systematic search strategies," in *Artificial Intelligence Planning Systems: Proceedings of the first international conference*, pp. 145–152, 1992.

[11] L. Davis, *Handbook of Genetic Algorithm.* Von Nostrand Reinhold, 1991.

[12] D. E. Goldberg, *Genetic Algorithm in Search, Optimization and Machine Learning.* Addison-Wesley Pub. Co., Inc., 1989.

[13] J. Holland, "Some practical aspects of adaptive systems theory," *Electronic Information Handling*, pp. 209 – 217, 1965.

[14] A. Bethke, "Genetic algorithms as function optimizers," Tech. Rep. 197, Logic of Computer Group, University of Michigan, USA, 1978.

[15] A. Chalmers and S. Gregory, "Constructing minimum path configurations for multiprocessor systems," Tech. Rep. CSTR-92-12, Computer Science Department, University of Bristol, UK, Apr. 1992.

[16] T. Warwick, *A GA Approach to Constraint Satisfaction Problems.* PhD thesis, Department of Computer Science, University of Essex, UK, 1995.

[17] T. L. Lau and E. P. K. Tsang, "Applying a mutation-based genetic algorithm to the processor configuration problem," in *IEEE 8th International Conference on Tools with Artificial Intelligence*, 1996.

[18] T. L. Lau and E. P. K. Tsang, "Solving the processor configuration problem with a mutation-based genetic algorithm," *International Journal on Artificial Intelligence Tools*, vol. 6, no. 4, 1997.

[19] C. Voudouris, *Guided Local Search.* PhD thesis, Department of Computer Science, University of Essex, UK, 1997.

[20] E. P. K. Tsang and C. Voudouris, "Fast local search and guided local search and their application to british telecom's workforce scheduling problem," *Operations Research Letters*, vol. 20, pp. 119–127, Mar. 1997.

[21] C. Voudouris and E. P. K. Tsang, "Guided local search," Tech. Rep. CSM-247, Department of Computer Science, University of Essex, UK, Aug. 1995.

[22] C. Voudouris and E. P. K. Tsang, "Function optimization using guided local search," Tech. Rep. CSM-249, Department of Computer Science, University of Essex, UK, Sept. 1995.

[23] C. Voudouris and E. P. K. Tsang, "Partial constraint satisfaction problems and guided local search," in *Practical Application of Constraint Technology*, pp. 337–356, Apr. 1996.

[24] S. Tiourine, C. Hurkens, and J. Lenstra, "An overview of algorithmic approaches to frequency assignment problems," in *CALMA Symposium, Scheveningen*, 1995.

A decision aid to predict monthly signal coverage maps between 30 and 50 GHz in Europe

Anil K. Shukla, Ammad Akram
Radio Science and Propagation Group,
Defence Evaluation Research Agency,
Malvern, Worcs., WR14 3PS.
United Kingdom.

Tad Konefal
Department of Electronics
University of York
Heslington
York, YO1 5DD
United Kingdom.

Peter Watson
Department of Electronics
University of Bath
Bath, BA2 7AY
United Kingdom.

1. Summary

The requirements of both commercial and military communications operators are towards increased data rates at long ranges to support applications such as the Integrated Digital Battlefield, High Definition Video Transmission (HDTV) and personnel communications for both micro and pico cell coverage. These requirements are forcing system designers and technology engineers to consider ever higher frequencies, such as millimetre wave frequencies (e.g., 20-40 GHz), where the spectrum is currently uncongested and where the necessary wide bandwidths may be available. As the number of commercial and military systems operating within this band increases, the pressure to optimise the packing density whilst minimising spectrum usage, interference and costs will also increase.

To aid the effective design and deployment of tri-service millimetre wave systems, a first generation **MI**llimetric **D**ecision **A**id **S**ystem (MIDAS Version-1.0a) has been developed. This aid predicts monthly signal attenuations and availabilities between (initially) 30-50 GHz in the European environment. The tool may be used by system designers for hardware development optimisation, and by system planners to match communications tactics (e.g., platform altitude, range) to the battlespace environment to maximise the operational effectiveness of limited and costly communications assets.

2. Introduction

The Integrated Digital Battlefield concept relies on the free flow of information between intelligence gathering sensors, battle management platforms and response teams in dynamic military engagements. The need to accurately characterise the battlespace and hence optimise the decision making process has lead to a continual increase in the number and complexity of sensors deployed. Consequently, the development of extremely high data rate (>150 Mbit/s) long range (>100 km) radio communication systems is receiving widespread attention within the military community.

The route to radically higher data throughput lies with increased system bandwidth which inevitably involves operating at higher frequencies in the electromagnetic spectrum. The bandwidth resources needed to support proposed extremely high data rate systems have to be sought at frequencies >10 GHz. Even at such high frequencies, however, the spectrum is congested by both military and commercial users. Interest, therefore, is growing in frequencies between 20-40 GHz, for both military and commercial applications, which aim to exploit the emerging semi-conductor technologies to develop cost efficient reliable systems.

Military systems exploiting these higher frequencies may also realise additional benefits such as the use of narrow beamwidth (~1-2 degrees) antennas to minimise interception and jamming, the opportunity for frequency re-use within a deployment, and the potential for exploiting millimetric technology to reduce communications payload and size.

Radiowave propagation through the Earth's atmosphere at millimetric wavelengths is affected by a number of tropospheric phenomena [*Watson and Brussard,* 1989]. Tropospheric constituents such as gases, clouds, water, humidity (i.e. hydrometers), tropospheric turbulence and stratified layers all contribute to time varying absorption, scattering, diffraction, refraction, reflections and depolarisation. The impact of these phenomena on

Paper presented at the RTO IST Symposium on "Frequency Assignment, Sharing and Conservation in Systems (Aerospace)", held in Aalborg, Denmark, 5-7 October 1998, and published in RTO MP-13.

millimetric systems may be predicted using decision aids.

Generally communications decision aids can be categorised as either engineering or tactical aids. The former are primarily used by systems engineers in considering the statistical nature of the propagation environment to determine the long term (e.g., annual) system performance. Tactical decision aids are most useful to military planners in considering propagation effects on systems in theatre and subsequent modification of military tactics such as flight plans and deployment ranges. Tactical decision aids, therefore, must account for a range of time scales from months, to facilitate planning, to hours to exploit the time varying environment.

Affordable and portable computing resources have led to tactical decision aids, such as HF-EEMS [*Shukla et. al*, 1996; *Lewis and Beattie*, 1996] and RPOT [*Hitney et. al.*, 1996], to become increasingly important tools for military planners and systems operators [*Cannon et. al.,* 1997]. The role and dependence of these tools in mission planning and deployment is predicted to increase. This will ensure optimum utilisation of limited and expensive assets in an environment where the ultimate system performance is constrained by the propagation channel.

This paper describes the monthly millimetre wave decision aid system (MIDAS Version 1.0a) being developed at DERA in collaboration with the University of Bath. The aid currently supports the design and development of long range (>100 km) airborne millimetre wave communications systems. The paper first introduces modelling and system techniques and then outlines the advantages of the MIDAS tool. Two key modules are then described and MIDAS results from an illustrative scenario are presented.

3. Propagation and systems overview

The dominant tropospheric phenomena impacting on radiowave propagation at millimetre wave frequencies are hydrometeors, especially in the form of precipitation (i.e. rain and clouds). Atmospheric gases comprising oxygen and water vapour (often termed collectively as the clear air effects) absorb radiowaves at certain characteristic frequencies and produce omni-present but slowly varying attenuation bands.

Water vapour presents three distinct absorption lines (see Figure 1) at 22.3, 183, 324 GHz while oxygen absorption lines are present at 60 and 119 GHz. The clear air attenuation decreases with increasing altitude as illustrated in Figure 1 by curve B. Frequencies near 60 GHz may be used for short range (< 1 km) tactical systems by exploiting the local peak in specific attenuation. Requirements for longer ranges systems, however, force system designers towards the local attenuation minimum near 27 GHz.

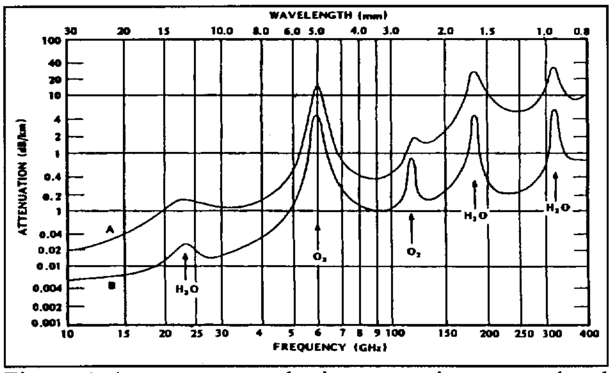


Figure 1 Average atmospheric attenuation at sea level (curve A, T= 20°C, P= 760 mm P_{H_2O} = 7.5 g/m^3) and at an altitude of 4 km (curve B, T= 0°C, P_{H_2O} = 1 g/m^3).

The slowly varying clear air losses can be predicted using models such as CCIR [*CCIR*, 1990] or FASCODE [*Anderson and Chetwynd,* 1992] with a high degree of reliability. Predicting the variable tropospheric losses (i.e. due to hydrometeors), however, can only be achieved using extensive meteorological databases. Scattering and absorption by hydrometeors such as rain, clouds and the melting layer are the most important mechanisms in determining fade margins for systems with availability up to 99% at millimetric frequencies. Rain attenuation can exceed several tens of dBs, for a typical path length of a few kilometres, during heavy downpours. For higher availability systems (>99.9%) additional tropospheric effects cannot be overlooked. These include: signal scintillation due to atmospheric turbulence, scattering and absorption by fog, refraction and reflection in tropospheric layers as well as depolarisation by rain and ice particles.

Current system design and development practices aim to allocate adequate fade margins to overcome the attenuating effects of the propagating environment and ensure a predetermined, cost effective, degree of signal availability. This method may be appropriate for fixed service links where a fade margin has been determined for a specific geographic location and the system designed accordingly. The attenuating tropospheric phenomena, however, vary on spatial scales of a few kilometres (e.g., localised rainfall) to many hundreds of kilometres (e.g., weather fronts) and exhibit latitudinal (e.g., tropical, arctic) and temporal variations of hours, seasons and even years. A mobile system, therefore, designed using annual fade margins may not meet operational requirements due to globally diverse climatic environments.

One technique to ensure successful operations in widely diverse environments is to design and develop systems with sufficient fade margins to overcome all potential environmental conditions. This, however, may prove to be prohibitively expensive, particularly for airborne platforms. The technique may also not be cost effective, particularly, if the conditions to be overcome only occur for a very small percentage of time and are coincident with military operations not taking take place (e.g., heavy thunderstorms). Antennas and amplifiers are amongst the most costly elements of mm-wave systems. If the antenna comprises a phased array, then a 3 dB fade margin increase is equivalent to doubling the number of array elements. This has obvious consequences for both cost and complexity of the overall system.

An alternative technique to high fade margin systems is to develop a system with a cost effective lower fade margin and deploy the system with a decision aid tool (such as MIDAS) which could match the planning tactics (e.g., link range, platform altitudes, data rate) to the battlespace environment. The technique may have the added advantage of reducing the constraints (i.e. availability) imposed on the system designer, enabling a more cost effective system to be developed.

A number of empirically based tools and procedures such as *CCIR* [1990], *Clough et. al.,* [1981], *Liebe* [1985], and *Crane*[1980] predict the long term (annual) attenuation due to the troposphere. The CCIR [*CCIR*, 1990] technique has the disadvantage that results are based on 0.01% annual rainfall data and other availabilities are empirically determined. It also suffers from the use of gross rain zone maps where physically improbable discontinuities occur between differing rain zones. The techniques of *Clough et al* [1981] and *Liebe* [1985] predict specific attenuations for specific user defined environments and provide only limited availability information. *Crane*[1980], however, is able to predict availabilities, but this technique is based on measured rain cell structures from rain gauges networks deployed up to ~22km in range, and may not be directly applicable to long range (>50 km) links. The available models were not, therefore, considered ideal for long range air mobile geometries where signals may traverse a number of tropospheric phenomena (e.g., rain and cloud and melting zone) and the platform may operate in differing climatic conditions.

To overcome current modelling limitations and provide system designers and planners with higher spatial and temporal resolution data, an improved tropospheric attenuation model has been developed by *Konefal and Watson* [1997]. The enhanced algorithms make use of the best currently available models and databases and is a key element of MIDAS.

4. Decision aid overview

The current version of MIDAS (Version 1.0a) is optimised for monthly low fade margin (less than 99% availability) terrestrial, airborne and satellite systems. The operational frequency range is 30-50 GHz and geographically restricted to the European environment (i.e. from 24°N, 27°W to 73.5°N, 45°E). Both of these aspects, however, have been enhanced [*Akram et al.,* 1999]. The decision aid comprises 5 key elements (Figure 2): input/update interface, tropospheric model, propagation prediction model, communications model and systems recommendations model.

The input/update interface is the human communications interface (HCI). The tropospheric model characterises the environment through which the mm-wave signals will propagate. Typical outputs from the tropospheric model such as water vapour density, rain rate, rain height, cloud location, are used by the propagation model. This latter model first determines the path length through various tropospheric features (e.g., path length through rain and clouds) and then predicts the attenuation and associated probability. The communications module uses the predicted values of attenuation and availability and would, ideally, compare them against a specified system performance and automatically suggest system trade-offs such as range and altitude to satisfy the users requirements. The output module is used to display the predictions in the most effective manner.

In MIDAS Version 1.0a the primary emphasis is on the tropospheric and propagation attenuation prediction models and are both described in more detail below. For point to point predictions, the communications module uses such information as antenna gains, system noise, bit-error-rate (BER) and modulation schemes. For signal coverage predictions excess attenuation are used at this early stage. (Excess attenuation is the variable attenuation, above fixed free space loss, due to the variable atmosphere that must be overcome and is a key element in fade margin calculations). The system recommendation model is currently limited to operator interpretation using the displays available.

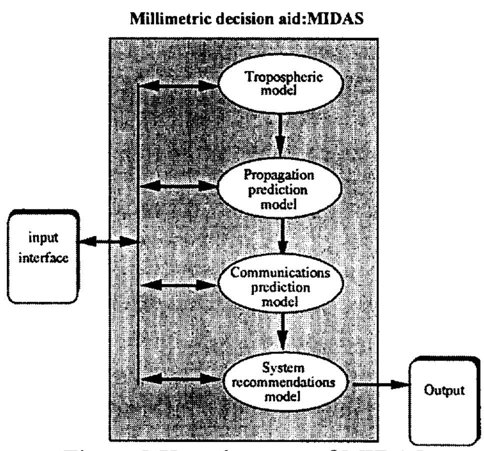


Figure 2 Key elements of MIDAS

4.1 Tropospheric model

The tropospheric components considered in the model are rain, melting layer, cloud, water vapour and oxygen. For each component a monthly probability of exceedance is derived with the predicted attenuation. This is not necessary in the case of oxygen due to its persistence and lack of variability in time. Each component is briefly outlined in (a)-(e) below.

a) The rain model. The method of *Tattelman and Scharr* [1983] is used to predict monthly 1 minute rainfall rates R_p at six exceedance levels, namely 0.01, 0.05, 0.1, 0.5, 1.0 and 2.0%. The rainfall rates are predicted from an equation of the form:

$$R_p = A_p + B_p T + C_p I + D_p f(L,T) \qquad (1)$$

where coefficients *A* to *D* are regression coefficients at exceedance level *p* and *f*(*L*,*T*) is a latitude temperature term [*Tattelman and Scharr* [1983]. *T* is the monthly mean temperature (^{0}F or 1.8x^{0}C+32) and *I* is the precipitation index (monthly precipitation/number of rainy days in mm per day).

The monthly data required to predict R_p are mean temperature, mean precipitation and number of rainy days. Grids of these parameters using *Leemans and Crammer*, [1991] data have been prepared for the European region using the MIN_CURVE_SURF function of IDL® [1995].

The 0^0C isotherm height required to derive path lengths through rain and other calculations is known to vary with season, month and latitude. Consequently, the mean monthly heights and corrections are predicted for the European region using the algorithms detailed in *Watson et. al.,* [1987].

b) The melting layer model. The melting layer, which is a mixture of ice-water and air and often observed as a radar bright band during stratified rain, is not explicitly modelled. Instead it is treated as rain which is assumed to extend to the 0^0C isotherm. *Russchenberg and Ligthart* [1996] indicate that in the frequency range 30-50 GHz this assumption will tend to overestimate the attenuation and will result in an underestimate of signal availability.

c) The cloud model. Cloud attenuation is predicted using the *Salonen and Uppala* [1991] model which uses, $W_{red}(p)$, the integrated reduced liquid water content of the cloud, as input. The model assumes that cloud liquid water is present between the 0^0C and -20^{0}C isotherms. If a temperature gradient of -5^{0}C/km with height is assumed, the cloud region is predicted to have thickness of $h_c \approx 4$ km.

The annual data for W_{red} published by *Salonen and Uppala* [1991] has an exceedance which fit an equation of the form $P(W_{red}) = \alpha e^{-\beta W_{red}}$, where α and β are local geographic regression coefficients. Assuming the same form of equation for the exceedance statistics of each month (i.e. $P_m = \alpha_m e^{-\beta_m W_{red}}$) it is readily deduced that $\beta_m = \beta$ and $\alpha = \overline{\alpha_m}$ where the bar denotes the mean value of α_m over 12 months.

The quantity $\alpha_m = P_m(W_{red} > 0) = P_m(Cloud)$, the probability of cloud being present for the month. It is assumed that this probability is proportional to the mean monthly humidity at the 0^0C isotherm, for which grids have been prepared for Europe using data supplied by ESA of mean radiosonde profiles over 10 years for 62 European sites. Denoting the mean monthly relative humidity at the 0^0C isotherm by U_m and writing $\alpha_m = kU_m$ we find :

$$\alpha_m = \alpha \frac{U_m}{\left\{ \frac{1}{12} \sum_{m=1}^{12} U_m \right\}} \qquad (2)$$

This has been used to derive monthly statistics for the exceedance of W_{red} required by *the Salonen and Uppala* [1991] attenuation model.

d) The water vapour model. The *Salonen and Uppala* [1991] annual water vapour database gives annual exceedance statistics for the integrated water vapour content *V*(kgm^{-2}) in the form $V(p) = a - b \ln p$ where *a* and *b* are local regression coefficients. The surface value of water vapour content ρ_s (gm^{-3}) is given by *Salonen and Uppala* [1991] as $\rho_s(p) = V(p) / h_s$ where h_s is the scale height, equal to 2.93km in rain conditions. Combining these models, the annual exceedance probability of water vapour (ρ_s) is given by:

$$P(\rho_s) = \alpha e^{-\beta \rho_s} \qquad (3)$$

$$\text{where } \alpha = \exp\left(\frac{a}{b}\right) \text{ and } \beta = \frac{h_s}{b}.$$

Using the same methodology as for deriving monthly cloud parameters, the monthly exceedance of ρ_s from $P_m(\rho_s) = \alpha_m e^{-\beta_m \rho_s}$ where $\beta_m = \beta$ and $\alpha = \overline{\alpha_m}$. We note that $\alpha_m = P_m(\rho_s > 0)$ and predict that this probability is proportional to ρ_m, the mean value of ρ_s during month m. The value of ρ_m can be found from the equation

$$\rho_m = \frac{216.5}{(273 + t_m)} U_m \times 6.11 \exp\left[\frac{19.7 t_m}{273 + t_m}\right] \text{gm}^{-3} \quad (4)$$

where t_m is the mean surface temperature (^{0}C) and U_m is the mean surface relative humidity ($0 \le U_m \le 1$). We write $\alpha_m = k'\rho_m$ and deduce

$$\alpha_m = \alpha \frac{\rho_m}{\left\{\frac{1}{12} \sum_{m=1}^{12} \rho_m\right\}} \quad (5)$$

ρ_s is thus associated with an exceedance probability p where

$$\rho_s(p) = \frac{1}{\beta} \ln \alpha_m - \frac{1}{\beta} \ln p \quad (6)$$

e) Oxygen model. Oxygen attenuation is calculated using a model atmosphere at 15^{0}C and 1013 ± 50mbar at the Earth's surface. The equivalent height for oxygen is taken to be $h_0 = 6$km [*CCIR*, 1990].

4.2 Propagation and attenuation model

The propagation model attempts to take into account the effect of refraction in the atmosphere by assuming a 4/3 physical Earth radius. The propagation model does not account for multipath, scattering, scintillation, or ducting effects. Depending on the precise platform co-ordinates, there may or may not be a path through tropospheric features. The propagation model first determines the path length through the differing components and then evaluates the attenuation and associated probability. The methods used for each tropospheric component are outlined below.

a) Attenuation due to rain. The rainfall rates are converted into attenuation using the method of *Leitao and Watson* [1986]. In this method the attenuation due to a point rainfall rate R (mm/hr) is given by

$$A_{rain} = a\left(\frac{R}{20}\right)^{x,y} L_S \text{ dB} \quad (7)$$

where L_S is the path through rain, in km, and the coefficients a, x and y depend on frequency and the horizontal projection L_g on to the earth of the path L_S. The choice of coefficient x or y depends on whether or not R is less than or greater than 20 km respectively.

The section of link path below the 0^{0}C isotherm defines L_S, the path through rain. To model both showery and widespread rain, two sets of coefficients for a, x and y have been defined by *Leitao and Watson* [1986]. In the model, showery rain coefficients are used at or below 0.1% exceedance and widespread coefficients above 0.1%. The coefficients are derived assuming circular polarisation.

b) Attenuation due to clouds. *Salonen and Uppala* [1991] give the annual zenith cloud attenuation as $A_{cl}(p) = gW_{red}(p)$ dB, where g is a known frequency dependent constant and On aircraft-to-aircraft links of distance, d_C, the attenuation will be approximately $g\frac{W_{red}}{h_c}d_C$ dB, for sufficiently small distances d_C. For distances much larger than 4km, an integration must be performed.

c) Attenuation due to water vapour. The specific attenuation of water vapour γ_w is predicted by [*CCIR*, 1990] as:

$$\gamma_w = \left\{ 0.05 + 0.0021\rho + \frac{3.6}{(f - 22.2)^2 + 8.5} + \frac{10.6}{(f - 183.3)^2 + 9.0} + \frac{8.9}{(f - 325.4)^2 + 26.3} \right\} f^2 \rho \times 10^{-4} \text{ dB/km} \quad (8)$$

where the frequency f is in GHz and water vapour content ρ is in gm^{-3}. The equation is valid at a temperature of 15^{0}C and pressure of 1013 ± 50mb. A correction of -0.6%/^{0}C is needed at temperatures higher than 15^{0}C (γ decreases slightly with increasing temperature). Equation (8) can be used to define monthly surface exceedance values for γ_{ws} using the value of $\rho_S(p)$ given by Equation (6). It is assumed that at a height h above the surface the corresponding specific attenuation is given by $\gamma_w(h,p) = \gamma_{ws}(p)\exp(-h / h_w)$ where γ_{ws} is the surface value and $h_w = 2.1$km in rain conditions [*Salonen et al.*, 1994]. Water vapour will be present along the entire link path, and the exponential factor will produce considerable local variations of $\gamma_w(h,p)$ along the path. This suggests that a numerical integration method should be employed.

The link path is divided into 500 sections, and the height h_q above the earth's surface for each section located at $\mathbf{r}_q$ is determined (q=0 to 499). A corresponding value of $\gamma_w(h_q, p)$ is computed at each point $\mathbf{r}_q$ in the following manner. A saturation temperature is computed for the surface value of $\rho_{sq}(p) = V_q(p) / h_s$, given by:

$$t_{satq} = 14 \ln(0.22\rho_{sq}(p))\,^{o}C \qquad (9)$$

If $t_{satq} \leq 15^0 C$, the value of ρ_{sq} is achievable at 15°C and the temperature correction factor $C_{tq} = 1$. If $t_{satq} > 15^0 C$, we need to reduce γ_{wsq} (calculated from Equation 19) by 0.6% per °C above 15°C, i.e.

$$C_{tq} = \left(1 - \frac{0.6(t_{satq} - 15)}{100}\right) \qquad (10)$$

The value of attenuation coefficient which is exceeded with probability p at the point $\mathbf{r}_q$ on the link path is then given by

$$\gamma_w(h_q, p) = \gamma_{wsq}(p) C_{tq} \exp(-h_q / h_w) \qquad (11)$$

An upper bound on the total water vapour attenuation for the link path is found by correlating the elements of attenuation produced from each of the 500 sections of the link path, adding attenuations for equal exceedances. Hence the water vapour attenuation statistics are given by

$$A_{wvap}(p) = \sum_{q=0}^{499} \gamma_w(h_q, p) \frac{|\mathbf{r}_2 - \mathbf{r}_1|}{500} \text{ dB} \qquad (12)$$

where $|\mathbf{r}_2 - \mathbf{r}_1|$ is the distance between the two end points of the link path and must be in units of km.

d) Attenuation due to oxygen. The surface specific attenuation of oxygen is given in these conditions as

$$\gamma_{os} = \left\{ 7.19 \times 10^{-3} + \frac{6.09}{f^2 + 0.227} + \frac{4.81}{(f - 57)^2 + 1.5} \right\} f^2 \times 10^{-3} \text{ dB/km} \qquad (13)$$

At a height h_q, corresponding to a position $\mathbf{r}_q$ on the link path, the attenuation coefficient is given by:

$$\gamma_{oq} = \gamma_{os} \exp(-h_q / h_o) \qquad (14)$$

The total oxygen attenuation on the link path is found by numerical integration:

$$A_{ox} = \sum_{q=0}^{499} \gamma_{oq} \frac{|\mathbf{r}_2 - \mathbf{r}_1|}{500} \text{ dB} \qquad (15)$$

with the link path $|\mathbf{r}_2 - \mathbf{r}_1|$ expressed in km.

In order to place an upper bound on the overall attenuation statistics, it is desirable to correlate the various factors described earlier, adding attenuations for equal exceedances. Correlating the attenuations in this way has a physical basis in that water vapour, cloud and rain are indeed associated with each other. For example, the formation of cloud requires sufficiently high water vapour content, and the formation of rain generally requires high cloud content. Using this technique of combining attenuations, and assuming the maximum number of terms in the equation (two cloud terms and one rain), we obtain

$$A_{Tot}(p) = A_{ox} + A_{wvap}(p) + A_{cl1}(p) + A_{cl2}(p) + A_{rain}(p) \qquad (16)$$

In cases where the aircraft-to-aircraft link path does not pass through rain, there will be a maximum of one cloud term and the term $A_{rain}(p)$ is dropped. At higher altitudes there may not even be a cloud term in which case only the terms in oxygen and water vapour remain. These two terms are present on every link.

4.3 Typical MIDAS inputs and outputs

MIDAS has three primary outputs. The first are simple signal coverage maps of either signal availabilities for a fixed excess attenuation (which dominates fade margin calculations) or excess attenuation for a fixed availability. Signal availability is defined as the percentage of time that the attenuation is less than some specified excess attenuation. The model inputs for this mode are: transmitter location in terms of altitude (km), latitude (°N), longitude (°E of Greenwich), receiver altitude (km), frequency (GHz), month (1-12), whether interested in availability or attenuation, desired availability or attenuation threshold and finally the boundaries of the region centred at the transmitter range (in km) bounding the mobile location of the receiver.

A typical signal coverage output is shown in Figure 3 for April (Month 4). In this example, transmitters are located over Glasgow and London at altitudes of 10 km and the receivers (at altitudes of 0.1 km) are assumed to be located within a range of ±200 km centred on the transmitter. The colour contours are signal availability, at 36 GHz, for an excess attenuation of 15 dB.

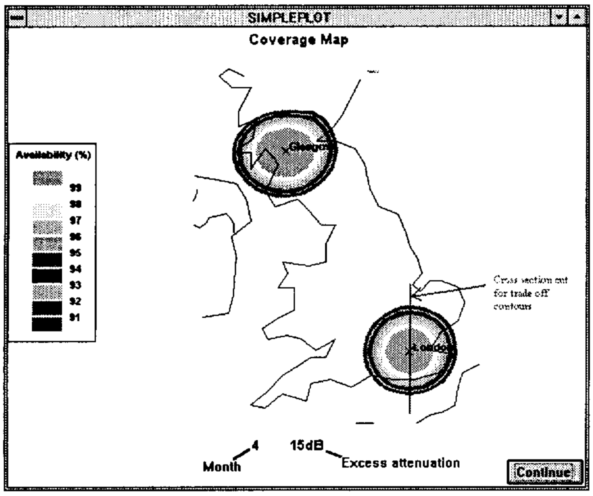


Figure 3. London and Glasgow coverage at 36 GHz

Figure 4, the second output format, shows link margin (i.e. excess attenuation dBs, abscissa) to be overcome, range (km, ordinate) and colour contours of signal availability. The range is along a line of longitutde centred on the transmitter location (in this case going due North of the London transmitter Figure 3). The plot is designed to enable system designers to trade-off between the three important parameters of: link range, availability and attenuation. For example, an attenuation of up to 18 dB will be observed less than 99% of the time to a range of ~100 km. At ~160 km, however, 18 dB is predicted less than ~93% of the time.

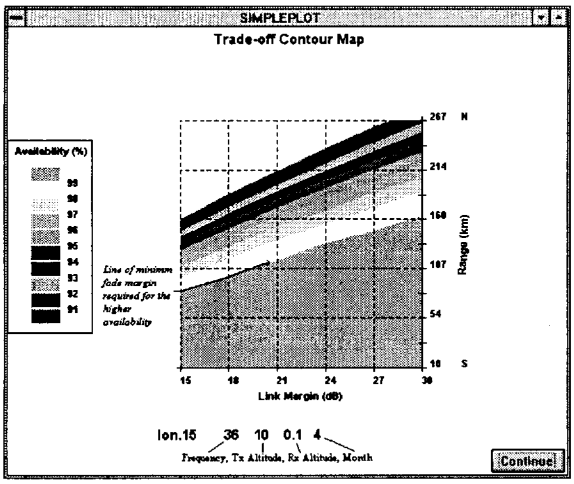


Figure 4 A system trade off contour plot

The final output is from a module which predicts point-to-point link budget analyses [*Akram et. al.*, 1996]. The output takes into account the system gains (e.g., transmitter power, antenna gain etc.) input by the user, and theoretical curves of typical modulation schemes to derive a monthly system performance. A typical input and output screen are illustrated in Figure 5.

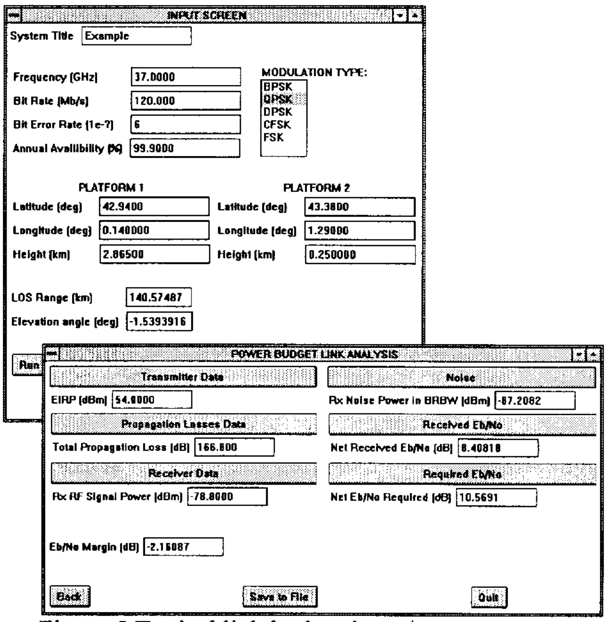


Figure 5 Typical link budget input/output screens

5. Application of MIDAS to an illustrative European scenario

An example scenario is now considered to illustrate the utility of the MIDAS tool. The scenario assumes that a 36 GHz system, developed to overcome a fixed excess attenuation of 15 dB, is to be deployed in London in December and July to provide a communications availability up to 99%, up to a range of ~90 km. Transmitter and receiver altitudes of 10 km and 0.1 km are initially assumed and signal coverage is predicted as a function of month. The results are used to estimate the number of additional transmitters required to achieve a fixed signal coverage within a large geographic area. The spatial variation is then predicted at a number of locations in Europe.

5.1 Monthly coverage variation over London

Rainfall can broadly be categorised as showery (often termed convective rain) and widespread (stratiform rain). The former, is associated with intense short duration rainfall events and is often observed in the summer months in a temperate climate. The latter exhibits lower rain rates of longer duration and is usually observed in the winter months. The attenuation, per unit distance, observed is greater for showery than widespread rain due to the increased rain rates.

Generally, London is within a temperate climate zone. It is predicted, therefore, that the reception range (for signal availabilities up to 99%) will decrease in summer than in winter. This is confirmed by trade-offs Figures 6a and b.

Figure 6a is the predicted plot (along a line of longitude centered on the transmitter location, Figure 3) for December whilst Figure 6b is for July. In winter (December) the 15 dB excess attenuation extends out in range (due North) to ~87 km for up to 99% of the time. In summer (July), however, the due North range decreases (by ~26%) to ~68 km. The requirements, therefore, are almost met in December but not July.

Figure 7 shows the monthly variation in range, due North and South, of the 15 dB excess attenuation. The predicted range due South is, on average, ~10 km shorter than that due North. This may be attributed to a combination of increasing melting layer height and, possibly a smaller increase in rain rates observed towards the South. The anisotropy, however, does suggest that to achieve an isotropic signal coverage higher radiated powers are required due South than due North. This may significantly impact on system design (e.g. antenna footprint) particularly on systems deployed near coastal environments where climatic conditions may be more variable.

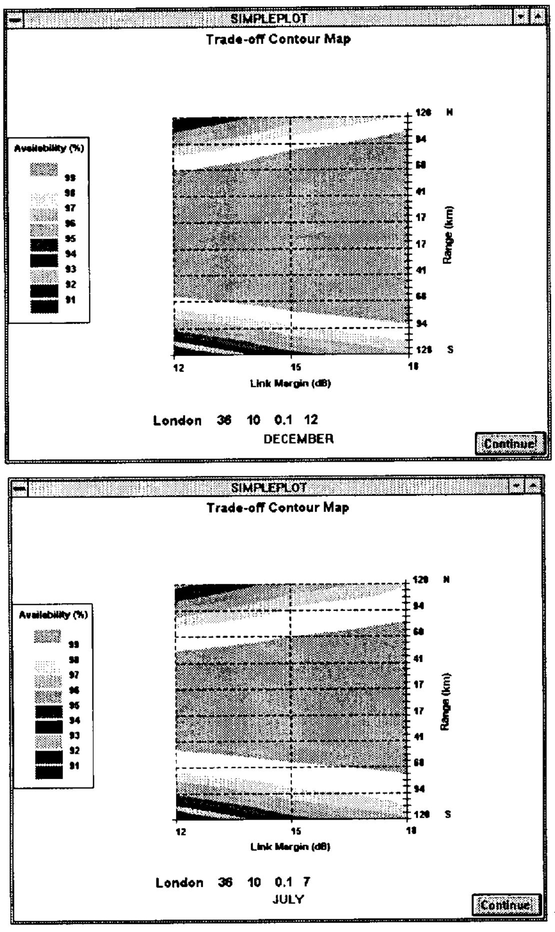


Figure 6 North-south, trade-off-plot for London in a) December and b) July

Figures 6 and 7 show that although the requirements may almost be met in December they cannot be achieved in July. Indeed the requirements may only be met for a short period of the year. The availability (up to 99%) shortfall in July may be overcome by either deploying an additional system in theatre, or increasing the radiated transmitter power from the single system via higher gain amplifiers, or antennas. Predictions show that for the latter option an additional ~5 dB (Figure 6b) of system gain. Increasing the radiated power or deploying an additional system, however, may be prohibitively expensive.

An alternative option is to optimise the transmitter geometry e.g. by increasing/decreasing the transmitter altitude to match the monthly climatic conditions.

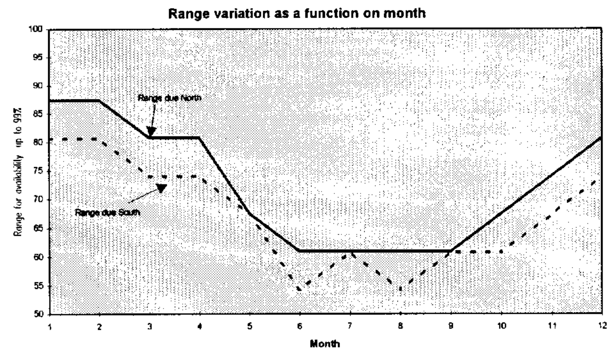


Figure 7. Monthly variation of range due North (solid line) and due South near London

Table 1 shows the impact on the range due North (for the 15 dB attenuation and availability up to 99% in July) of increasing the transmitter altitude from 10 - 13 km. As the altitude increases the range achievable increases, as expected, due to the reduced slant paths through the different tropospheric features. At an altitude of 13 km a range of ~87 km, which almost matches the scenario requirements, is predicted.

Table 1 Increasing antenna altitude in July

Antenna altitude (km)	Range (North)
10	~68
11	~74
12	~81
13	~87

The converse may also be achieved as shows in Table 2. As the altitude is decreased in winter the maximum range due North also decrease. At an altitude of 8 km a range of ~60 km, similar to that achieved in July, is predicted. This may be used to advantage if the coverage area is to be minimised to reduce interference.

Table 2 Decreasing antenna altitude in December

Antenna altitude (km)	Range (North)
10	~87
9	~74
8	~60
7	~54

5.2 Impact on packing density

In any planning phase of a wide area (e.g., 500 km) communications system deployment, the number of systems operated is usually minimised to ensure cost effectiveness whilst maintaining a pre-determined system availability. To illustrate the monthly variability in the number of systems required to maintain a specified coverage, a simplified scenario is considered below. In this scenario it is assumed that identical total coverage is required in December and July.

For simplification purposes symmetry is assumed and four identical systems are deployed within a 360 km^2 region. Although the packing configuration described by Figure 8 is not optimal, it is useful for illustrative purposes. The

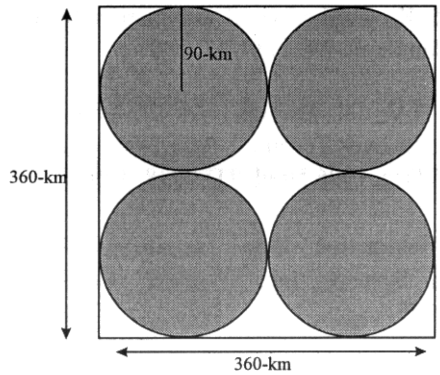


Figure 8 A simplified packing density model

In December the range due North with transmitters at an altitude of 10 km over London is ~90 km. Assuming symmetry the total percentage coverage from four transmitters within the 360 km square box-area is ~80%. In July, however, the range due North decreases to ~68 km and the total percentage coverage from four units decreases to ~44%. To ensure signal coverage similar to that of December, therefore, it is predicted that at least another 3 transmit stations would be required (totalling 7 within the area). Alternatively, the platform altitudes may be increased from 10 km to 13 km in July, to provide the increased coverage, significantly reducing the number of units required.

5.2 Coverage variation in Europe

In future military conflicts, systems may need to be deployed in diverse climates at short notice. A successful deployment scenario for the UK, however, may not be appropriate to other climates and system trade-offs may need to be made.

MIDAS predictions at a number of locations across Europe have been made to illustrate the spatial variation in tropospheric attenuation. It is again assumed that the deployment is to take place in December and that a signal availability up to 99%, to a 90 km range due North is required.

The locations investigated (Figure 9) are chosen for their differing climatic conditions as defined in *CCIR*[1990]. Figure 10 details the predicted range difference (due North) from London.

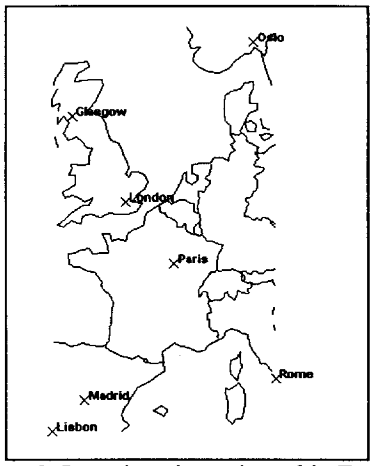


Figure 9. Locations investigated in Europe

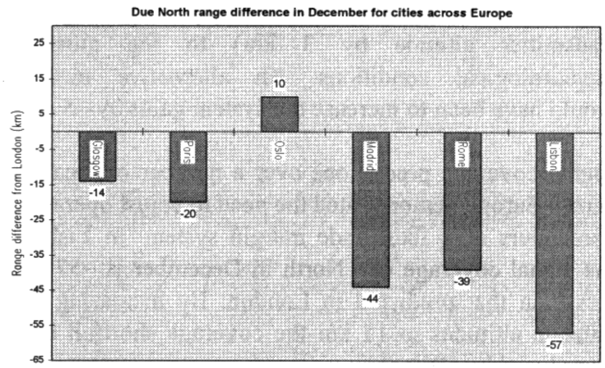


Figure 10 Range variation in Europe for December

The maximum range difference is between Lisbon and London where a difference of -57 km is predicted. To overcome this range shortfall additional system gain may be required to fill the availability gap or a degraded system performance accepted with the knowledge that the service will improve over the next six months. Alternatively, the environment can be mitigated against, to some degree, by increasing the antenna altitude to 15 km, where the shortfall is predicted to be ~27 km, a 50% improvement.

6. Conclusions

Decision aid tools are becoming increasingly important for design, planning and deployment of mobile communications assets which must function in diverse battlespace environments. At millimetre wave frequencies, the impact of hydrometeors predominates and tools which account for climate changes ranging in timescales, ideally, from minutes to years are becoming increasingly important. These decision aid tools may be used to mitigate against the detrimental effect of the battlespace environment which are difficult, or not cost effective, to overcome during the system design phase.

A 30-50 GHz monthly prediction model has been developed which calculates signal attenuation and availability due to rain, clouds, water vapour (in rain and non-rain conditions) and oxygen. The model is a key element of a **MI**llimetric **D**ecision **A**id **S**ystem (MIDAS Version 1.0a) being developed at DERA. MIDAS is currently used to provide:

(i) coverage maps of availability and attenuation
(ii) link margin-range trade-off maps
(iii) link budget analyses

The utility of the tool has been demonstrated using hypothetical scenarios in Europe. The first scenario assumed that a fixed 15 dB fade margin system was to be deployed in December, from a platform, at an altitude of 10 km, to provide availability coverage up to 99% up to a range of ~90 km due North from London. The tool was then used to overcome the ~20 km range variation predicted between December and July, by matching the operating scenario (by increasing the transmitter altitude by 3 km) to the monthly meteorological conditions. An alternative method would have been to increase the system gains by ~5 dB.

Signal coverage predictions over a number of location across Europe demonstrated the need to adjust operating parameters of a fixed fade margin system. In Lisbon, the signal coverage due North in December is ~57 km less than that predicted in London. By increasing the antenna altitudes to 15 km the coverage shortfall may be reduced by 50%.

A simplified packing model for a 380 km^2 region near London showed that to maintain a signal coverage of ~80% the number of systems deployed could range from four, in December, to seven in July. By carefully managing four assets, however, and optimising their deployment to the prevailing monthly climatic conditions, using a decision aid such as MIDAS signal coverage could be maintained whilst minimising the number of systems deployed in theatre.

7. References

Akram A., A.K. Shukla., T. Konefal, P. A. Watson, "A 10 - 50 GHz global predictions model" ICAP 1999 to be published.

Anderson G.P., J. H. Chetwynd, "FASCODE: Fast Atmospheric code preliminary version 3P" Hanscome Airforce Base Geophysical Laboratories Massachusetts, 1992.

Brussaard G. and P. A. Watson "Atmospheric modelling and millimetre-wave propagation". Chapman and Hall, 2-6 Boundary Row, London SE1-8HN, UK.

Cannon P. S., J. H., Richter and P. A. Kossey"Real-time specification of the battlespace environment and its effects on RF military systems", AGARD conference on Future Aerospace Technology in the Services of the Alliance, 14-16 April, 1997

CCIR "Recommendations and reports of the 17th general assembly", Report 719-3, CCIR Study Group 5, Radiometeorological data, Dusseldorf, 1990

Clough, S. A., F. X. Kneizys, L. S. Rothman and W. O. Gallery, Atmospheric spectral transmittance and radiance: FASCOD1B, Proc. Soc. Photo Opt. Instrum. Eng., 277,152-166, 1981.

Crane R. K., Prediction of attenuation by rain. IEEE Trans. Commun., 29, 1717-33, 1980.

Hitney, H.V., A. Barrios and W. Patterson "Radio Propagation over Terrain", Proc Battlespace Atmosph. Conf 3-5 Dec, 1996 Nrad, TD 2938, 1996

IDL® "Interactive Data Language Version 4.0". Research Systems Inc., Boulder, Colorado, USA., March 1995.

Konefal T., and Watson P. A., "Prediction of availability on mm-wave aircraft-to-aircraft radio links", DERA, final report, Contract CSM/584/UA, June 1997.

Leemans, R. and Cramer, W. "The IIASA database for mean monthly values of temperature, precipitation and cloudiness of a global terrestrial grid". International Institute for Applied Systems Analysis (IIASA). RR-91-18, 1991.

Leitao M. J. and Watson P. A. "Methods for prediction of attenuation on earth-space links based on radar measurements of the physical structure of rainfall". Proc. IEE F, 133(4) 429-40, 1986.

Lewis D., P. Beattie "The Electromagnetic Modelling system (EEMS) Project Overview" Radio Propagation over Terrain", Proc Battlespace Atmosph. Conf 3-5 Dec, 1996 Nrad, TD 2938, 1996

Liebe, H. J., An updated model for millimetric wave propagation in moist air, Rad. Sci., 20(5), 1069-1089, 1985.

Russchenberg, H.W.J. and Ligthart, L.P., "Backscattering by and propagation through the melting layer of precipitation: A new polarimetric model", IEEE Trans. on Geoscience and Remote Sensing, Vol. 34, No. 1, January 1996, pp3-14

Salonen E., and Uppala S., "New prediction method of cloud attenuation", Electronics Letters, Vol 27, No 12, , pp- 1106-1108, June 1991

Salonen E., Karhu S., Uppala S., Hyvonen R., "Study of improved propagation predictions" Final report for ESA, ESTEC contract number 9455/91/NL/LC(SC) Dec 1994.

Shukla A. P. S. Cannon., M Moore., "A HF tactical decision aid for conventional and automated Radio Control Systems", Proc Battlespace Atmosph. Conf 3-5 Dec, 1996 Nrad, TD 2938, 1996

Tattelman, P. and Scharr, K.G., "A model for estimating one minute rainfall rates", Journal of Climate and Applied Meteorology, Vol. 22, September 1983, pp1575-1580

Watson, P.A., Leitao, M.J. et al., "Prediction of attenuation on satellite-Earth links in the European region", IEE Proceedings, Vol.134, Pt. F, No.6, October 1987, pp583-596.

PAPER TITLE : **Paper 15 – A Decision-Aid to Predict Monthly Signal Coverage Maps between 30 and 50 GHzi Europe**

AUTHOR : A. Shukla

NAME : K.S. Kho

QUESTION :

1. Re. your fig. 7. How much should I increase the output power in order to have a coverage better than 90 Km for the whole year.

ANSWER :

To achieve the 90Km due north coverage you would need an additional ~5db of system gain to ensure coverage for the whole year. This may, however, be prohibitively expensive and thus not cost effective.

CONDITIONS D'INSERTION DES RADARS A LARGE BANDE

Jean ISNARD
THOMSON-CSF RCM
La Clef de Saint Pierre
1, boulevard Jean Moulin
78852 ELANCOURT
France

RESUME

La mise en oeuvre de signaux à large bande dans les radars résulte des besoins en discrimination / classification des cibles, résistance aux contre-mesures (passives et actives) et cartographie précises de certaines zones de terrain.

Suivant les applications les formes de signal émises ainsi que leurs conditions d'émission spatio-temporelles devront pouvoir être harmonisées entre elles et avec celles utilisées par des applications différentes (faisceaux hertziens, radiocommunications, ...) car les bandes de fréquences devront de plus en plus être partagées pour une meilleure utilisation de la ressource spectrale. Sont esquissées dans cet exposé les conditions internes et externes au Service de Radiolocalisation (RL) : le fonctionnement des radars à large bande avec d'autres services, notamment celui de radiocommunications, devrait être possible dans des bandes partagées grâce à l'acceptation de contraintes mutuelles.

SUMMARY

The implementation of wide-band signals in radars results of the need to discriminate / classify targets, to resist countermeasures and to obtain precise maps of certain areas.

Depending on the applications the signal waveforms and their spatial-temporal conditions of transmission will have to be harmonized among themselves and with those utilised by other applications; indeed it will become increasingly necessary to share at lest partly the bandwidths for a better use of the spectral ressource.

This paper outlines the internal and external conditions for the radiolocalisation service : the sharing of bandwidths between radars and other equipment in particular for radiocommunications should be possible provided that mutual contrains are accepted.

1. INTRODUCTION

De manière générale, les avantages attendus des radars à très large bande sont :

1) une discrimination / classification accrue des cibles,
2) une réduction des effets de glint et de multitrajet.

Une conséquence avantageuse résultant de la nécessaire répartition spectrale de l'énergie est :

3) une discrétion par dilution des signaux vis-à-vis des analyseurs spectraux

Dans le domaine des fréquences basses, on tire parti de la combinaison avec les propriétés de ces bandes [1]pour obtenir,

1) une visibilité accrue des cibles masquées (masques de végétation, pénétration du sol),
2) un processus de détection résistant mieux aux contre-mesures actives.

La très large bande peut être obtenue soit directement par l'émission d'un signal très bref (1 ns, par exemple), soit indirectement, à un degré moindre, par modulation du signal émis (type CDMA par exemple) ou par émissions étroites réparties séquentiellement dans une très large bande. Stricto sensu, seuls les radars du premier type ont droit à l'appellation très large bande (UWB) [2].

Une émission très large bande peut être définie par un signal dont le spectre Δf est tel que

$$\Delta f \geq 0,25 \frac{f_{max} + f_{min}}{2}$$

f_{max} et f_{min} représentant respectivement les fréquences maximum et minimum utilisées. On perçoit qu'il y a là un côté arbitraire, voire subjectif, que nous ne commenterons pas. Par la suite, nous prendrons aussi en considération les radars à large bande c'est-à-dire ceux dans la bande "instantanée" est supérieure ou égale à 10 % de la porteuse notamment dans les bandes "classiques" (L, S, C, X).

2. EXEMPLES D'APPLICATION

Afin d'obtenir une très haute résolution, la très large bande est recherchée dans de nombreuses applications [3] :

- exploration du sous-sol
- détection de mines
- détection de missiles furtifs anti-navires
- radar multimode aéroporté
- etc ...

Par exemple, l'apport d'une très haute résolution en radar aéroporté, dans les bandes habituelles C/X/Ku est notable dans tous les modes de fonctionnement dès que la bande dépasse 1 GHz.

1) en mode veille, la probabilité de détection croît en raison de la réduction du fouillis,
2) en modes suivis de terrain et de localisation, la haute résolution permet de discriminer les échos fixes tels que les pylônes par élimination des effets de glint et de séparer les cibles de leurs images (en pratique une résolution d'au moins 15 cm correspondant à une bande d'un gigahertz),
3) en mode reconnaissance / identification, la classification non coopérative des cibles (NCTR radar) apporte une contribution indispensable pour effectuer une identification ami-ennemi par fusion des divers moyens (data link, RWR, IFF, NCTR optro) notamment en vue de réduire le risque de tirs fratricides,
4) enfin la distance d'interception du radar est naturellement réduite par le fait que le récepteur d'écoute doit élargir ses bandes.

Paper presented at the RTO IST Symposium on "Frequency Assignment, Sharing and Conservation in Systems (Aerospace)", held in Aalborg, Denmark, 5-7 October 1998, and published in RTO MP-13.

3. SPECIFICITES LIEES A L'INSERTION SPECTRALE

L'objectif essentiel étant d'améliorer la résolution du radar pour extraire la réponse la plus détaillée possible de la cible permettant de la classer, voire de l'identifier, l'observation doit donc mesurer cette réponse le plus fidèlement possible vis-à-vis de la réponse théorique.

Pour faire cette mesure, le radar doit surmonter trois sortes de difficultés :

1) obtenir le meilleur pouvoir séparateur possible vis-à-vis de l'environnement (autres cibles ou brouillage),
2) déconvoluer les effets des fonctions de transfert des sous-ensembles du radar (émetteur - antennes - récepteur),
3) compenser les effets de propagation.

3.1 Pouvoir séparateur

On sait que pour obtenir à la fois une haute résolution en distance et en vitesse, il faut émettre un signal de produit BT élevé ; si on veut bénéficier des conditions simplificatrices de l'hypothèse dans laquelle l'effet Doppler est négligeable sur la durée de l'impulsion, le signal doit être limité à l'observation des cibles dont la vitesse radiale v satisfait à

$$BT \ll \frac{c-v}{2v}$$

Cette contrainte est encore satisfaite même pour des produits BT élevés.

Exemple :

$v \leq 1500$ m/s
$BT \ll 10^5$

Or pour un radar à $B = 3$ GHz
$T = 3{,}3$ µs
$BT = 10^4$

On sait par ailleurs que même la fonction d'ambiguïté (F.A.) généralisée possède des propriétés de continuité vis-à-vis de la F.A de woodward [4][5].

Le danger ne vient donc pas d'une détérioration rapide des propriétés de la F.A. mais du fait qu'en affinant la résolution distance, on ne collecte une quantité considérable de signaux provenant de sources parasites soit directement, soit indirectement via le clutter ou même la cible elle-même qui, dans le meilleur des cas, augmente la fausse alarme et dans le pire des cas empêche la mesure de la SER. : ce que l'on peut gagner en degré de liberté en haute résolution peut se perdre en interférence vis-à-vis d'autres cibles ou brouilleurs.

3.2 Déconvolution des fonctions de transfert des sous-ensembles radar

Par le fait même que chaque sous-ensemble Emetteur-Récepteur des radars a sa propre fonction de transfert, la forme d'onde idéale devant solliciter la cible doit être élaborée en conséquence.

On sait, par exemple, que le signal en sortie d'un émetteur connecté à une antenne dont la dimension est courte vis-à-vis de la plus petite longueur d'onde contenue dans ce signal est approximativement la dérivée du signal en entrée [6].

Il faut aussi reconsidérer la façon dont est spécifiée une antenne en très large bande car ses propriétés ne peuvent plus être définies autour d'un point moyen de fonctionnement stationnaire, au point que pour rendre optimum la liaison Emission-Antennes-Réception, il faille avoir des antennes différentes à l'émission et à la réception.

3.3 Propagation

Même si dans la partie 1 - 10 GHz du spectre les effets proprement dits de la propagation sont relativement stationnaires les interactions avec la matière notamment en pénétration à fréquences basses ou en réflexion (fouillis) sont très importantes pour les performances et les interférences.

De manière générale, les effets des interférences radioélectriques et des produits d'intermodulation dans des systèmes à grande dynamique comme les radars à très large bande sont un des problèmes majeurs de la compatibilité électromagnétique de ces systèmes.

4. ETALEMENT DE L'ENERGIE DANS LE PLAN TEMPS FREQUENCE

Certes les radars à très large bande sont un cas extrémal, mais la tendance générale des radars a été d'occuper de plus en plus l'espace temps x fréquence par la mise en oeuvre de techniques de compression d'impulsion, d'agilité de fréquence, de synthèse d'impulsion, etc ...

Les techniques de compression d'impulsion sont utilisées d'autant plus qu'elles sont indispensables à l'emploi des semi-conducteurs à faible puissance crête et des TOP à large bande en mode impulsionnel (facteur de forme typique de 30% et 10% respectivement).

Mis à part le cas encore peu répandu des radars à émission continue ou semi-continu, l'axe temporel du plan temps x fréquence est donc occupé par l'émission dans une proportion de 10 et 30% y compris les techniques de lever d'ambiguïté distance par variation de la période de répétition.

L'axe fréquence est occupé en proportion de la résolution, de l'agilité en fréquence, de la synthèse d'impulsion, ... De manière générale, la bande de fonctionnement représente environ 10 % de la porteuse, le radar cherchant à couvrir toute la bande attribuée et la bande du signal représentant le motif de durée minimale servant à extraire un plot correspondant à 1 ou 2 % de cette bande de fonctionnement.

Bande	S	C	X	
Fréquence centrale	3	6	10	GHz
Bande de fonctionnement	300	600	1000	MHz
Bande du motif. de signal	3	6/10	10/20	MHz

Vis-à-vis des bandes attribuées au service* de radiolocalisation (RL), la bande occupée par le motif minimum composant le signal ne représente que 1 à 2 % de cette bande.

On voit que l'introduction de radars à très large bande multiplie par environ 100 le champ d'occupation possible sur l'axe fréquence, la situation sur l'axe temps étant pratiquement inchangée, le récepteur du radar devant de toutes façons analyser les échos.

Or, cet étalement de l'énergie dans le spectre, et, dans une moindre mesure dans le temps, dépend de la forme du signal choisie pour l'application et de ses conditions d'émission spatio-temporelles. Ceci est illustré ci-après par quelques exemples :

1) les modes à très large bande sont très rarement utilisés pour la détection mais plutôt pour la classification : la séquence d'identification est généralement brève,
2) si la classification est faite en fréquences basses (100 MHz - 1 GHz) en sollicitant des résonances, le spectre est analysé séquentiellement,
3) lorsque le signal émis est élaboré par synthèse, celle-ci est faite à partir d'un certain nombre de raies discrètes,
4) les radars transhorizons ont une très large bande de fonctionnement possible (3-30 MHz) pour pouvoir adapter la fréquence d'émission aux conditions de propagation de l'instant,
5) dans les cas où l'identification doit se faire rapidement par des impulsions très brèves à spectre continu, ces impulsions sont si brèves que, dans la plupart des cas, elle ne perturbent pas de manière significative le fonctionnement des autres services.

5. ATTRIBUTION DES BANDES DE FREQUENCES (RR de l'UIT)

Dans le RR de l'UIT, le tableau d'attribution des bandes de fréquences au service RL fournit les informations.

On rappelle simplement ici qu'entre 1 et 36 GHz, il y a 8 secteurs principaux attribués mondialement au radar à titre primaire.

Certes certaines attributions sont à titre secondaire ou sont géographiquement limitées. Ce n'est pas l'objet ici d'entrer dans le détail des notes du RR ; il s'agit plutôt vis-à-vis du problème de l'insertion des radars à large bande dans le spectre d'avoir présent à l'esprit les secteurs attribués au radar et leur évolution probable.

* Le terme de service est pris ici au sens du RR Partie A Chapitre I, article 1 Section III. Exemples :
Service de radiolocalisation : service de radiorepérage aux fins de la radiolocalisation (radars)
Service de radionavigation aéronautique : service de radionavigation pour les besoins des aéronefs et la sécurité de leur exploitation
Service fixe : service de radiocommunication entre points fixes déterminés (faisceaux hertziens)

Ces secteurs sont en MHz

f_{min}	f_{max}	$0,25 \frac{f_{max} + f_{min}}{2}$
1215	1400	327
2700	3400	**762**
5250	5850	1387
8500	10 500	**2375**
13 250	14 300	3444
15 700	17 300	4125
24 050	26 250	6287
33 400	36 000	8675

Grosso modo, il n'y a que deux bandes qui puissent à peine admettre le fonctionnement de radars à très large bande suivant la définition rappelée au début ! Encore cette évaluation ne tient-elle pas compte de plusieurs convoitises sur la bande S.

Il existe tout un ensemble de normes et recommandations visant à obtenir la compatibilité électromagnétique interne et externe aux services ; par exemple, les émissions sont assujetties aux diverses dispositions du RR notamment à l'article 5 [7] :

- tolérance de fréquence (appendice 7)
- niveaux de puissance maximaux tolérés des rayonnements non essentiels (appendice 8)
- niveaux de puissance maximaux tolérés pour les émissions hors bande (appendices 17 et 27 Aer2)

En revanche, l'UIT recommande l'emploi des techniques d'étalement du spectre et les méthodes de traitement des signaux qui permettent d'utiliser le spectre avec le maximum d'efficacité.

Evidemment, la formulation de ces dispositions est faite dans une perspective "bande étroite" du point de vue radar : un réexamen de ces documents est à faire.

6. RECHERCHES GENERALES EN VUE D'UNE UTILISATION EFFICACE DU SPECTRE

Les besoins en spectre sont en augmentation pour toutes les applications dans des bandes déjà très utilisées, exemples : 1 à 3 GHz, 4 à 6 GHz, 8 à 12 GHz et même 30 à 36 GHz.

Deux axes de recherche sont les plus souvent cités [8] :

1) pour chaque service, il s'agira d'augmenter l'efficacité spectrale et la réutilisation des fréquences,
2) entre les services, il faut accroître le partage des bandes.

Il s'y ajoute le réaménagement du spectre, généralement en déplaçant nombre d' applications de radiocommunications vers des bandes plus élevées ; mais l'importance des investissements et la durée de vie de certains équipements (30 ans pour les radars par exemple) font que les effets bénéfiques n'en seront perçus qu'à très long terme.

De manière interne à chaque service, les techniques de modulation et de codage et la maîtrise des antennes sont probablement les sujets principaux ; du point de vue externe entre services, l'estimation des densités de puissance ou de puissance surfacique rayonnées est le thème majeur ce qui à nouveau impliquent les diagrammes de rayonnement des antennes ; mais on pourrait aussi y ajouter les contraintes d'exploitation dans la mesure où une coordination, voire une synchronisation est possible. Il s'agit avant tout de trouver une efficacité globale élevée, c'est-à-dire sensibilité aux brouillages moindre et capacité à réutiliser des fréquences : l'optimum interne peut devoir être dégradé au profit d'une efficacité globale supérieure.

Enfin l'évolution à court et long terme de l'utilisation de la ressource spectrale dépend de la connaissance des conditions de propagation des ondes dans les environnements réels, d'où le développement indispensable de méthodes et de modèles fondés sur des mesures.

7. CONDITIONS D'INSERTION

Evidemment les conditions d'insertion sont différentes suivant que l'on se place d'un point de vue interne ou d'un point de vue externe.

7.1 A l'intérieur du service RL

Les radars devant réserver la majeure partie du temps à capter les échos de cibles atteindront une bonne utilisation de la ressource spectrale sous deux conditions :

1) Découplage géographique,
2) Mise en oeuvre de plans de formes d'onde et de fréquences, voire de synchronisation d'émission.

Ce sont des conditions déjà mises en oeuvre dans le domaine Défense.

7.2 A l'extérieur du service RL

De ce point de vue, le radar a deux avantages :

1) Ses émissions sont entièrement synchronisées et de faible durée comparées à celles des autres services de radiocommunication,
2) La structure des signaux radioélectriques de la plupart des autres émissions lui sont connues puisqu'elles font l'objet de normes internationalement acceptées.

En revanche ses récepteurs doivent avoir une large bande avec une sélectivité bien inférieure à celle des radiocommunications par exemple.

Sous réserve de ne pas être saturantes, **les émissions radar** étant brèves, l'impact sur des liaisons de radiocommunication sera faible car celles-ci sont déjà concues :

1) pour résister aux effets des trajets multiples, par exemple la durée moyenne des évanouissements dus au fading de Rayleigh est de l'ordre de quelques ms dans la bande 150 à 900 MHz croissant en fonction de la longueur d'onde,
2) pour protéger les données contre les erreurs de transmission : des mécanismes de protection sont mis en oeuvre (entrelacement, codage convolutionnel, codage en bloc) ; sur un débit binaire de 22,8 kbits/s en GSM, 9,8 kbits/s sont utilisés pour la protection contre les erreurs.

A ce propos, les concepteurs de matériels d'infrastructures radiomobiles devraient ne pas trop réduire ces codages de protection pour tenter d'accroître le débit des canaux, car dans l'extension des bandes qu'ils souhaitent, ils en auront peut être besoin pour résister aux émissions des autres services dont le RL.

A la réception, le radar est confronté à des difficultés tout à fait analogues à celles des Moyens de Renseignement Electroniques (ELINT) pour localiser et éliminer les effets des brouilleurs représentés par les émetteurs des autres services.

La difficulté à séparer ces émetteurs est due :

- à l'observation de ces émetteurs à des distances de plus en plus grandes induisant une zone géographique instantanée très étendue,
- à la densité des émetteurs dans cette zone.

On dispose actuellement de plusieurs méthodes d'analyse temps x fréquence et corrélation spectrale (cyclostationnarité des signaux de radiocommunications) ainsi que d'algorithmes performants.

Les séparateurs aveugles, dits aux ordres supérieurs, exploitent en plus des informations dans les statistiques d'ordre deux des observations, celles contenues dans les statistiques d'ordre supérieur à deux, fondés sur l'hypothèse d'indépendance statistique des sources non-gaussiennes représentant les radiocommunications ; ces séparateurs ont déjà fourni des résultats très intéressants pour une complexité de calcul réalisable et sans nécessité de goniométrie préalable des sources.

Cette situation ne peut qu'inciter à une intégration future du radar et des contre-mesures au moins en ce qui concerne certaines fonctions ; en effet, dans les deux types de capteurs, on doit exploiter, de manière cohérente, des informations de nature spectrale recueillies sur des voies parallèles fournissant la discrimination spatiale ; même si les deux types ont des propriétés a priori différentes en termes de débits, de bandes d'analyse, de dynamiques, etc ... cette voie conduit à une architecture d'ensemble depuis les antennes jusqu'aux traitements.

On trouve un exemple de cette tendance d'associer les fonctions radar et mesures (ELINT) aux références [9][10] mises en oeuvre dans le cadre des radars CARABAS et LARISSA.

Bien entendu, une telle évolution ne pourra se faire que grâce au développement :

- de réseaux d'antennes à large bande,
- de modules de sélection de bande à large bande totale,
- de chaînes de réception à large bande,

- de processeurs de signal à calcul parallèle mettant en oeuvre des traitements de même type mais à paramètres différents.

8. CONCLUSION

C'est par un mutuel effort de compréhension que seront trouvées les solutions de fonctionnement des radars à large bande et d'autres équipements dans des bandes partagées.

Les radiocommunications doivent prendre en compte les spécificités du radar notamment en ce qui concerne les durées de réception et la dynamique des signaux. Les radars ont à mieux connaître les structures (temporelles, spatiales et de codage) des radiocommunications.

A long terme, un réaménagement du spectre sera nécessaire pour se débarasser aussi bien d'un découpage abusif du spectre en bandes trop étroites ne permettant aucune optimisation, que de rentes de situation telles que la radiodiffusion occupant en UHF (470-960 MHz) la moitié du spectre inférieur à 1 GHz alors que cette diffusion pourrait se faire par satellite. Du point de vue technique, il existe déjà de nombreux points de convergence entre radar et radiocommunications dans les domaines de la propagation, des antennes et du traitement de signal.

L'exemple de la propagation est instructif à ce sujet.

En radar, quelle que soit la bande mais particulièrement en bandes basses, la détermination des performances rend nécessaire l'élaboration de modèles prenant en compte les trajets multiples et les fouillis, notamment à incidences rasantes ; ceci a conduit à adopter une approche physique fondées sur les propriétés du champ électromagnétique en présence du sol et des obstacles divers : les meilleurs modèles pour des environnements inhomogènes semblent bien être ceux fondés sur des techniques hybrides d'équation parabolique et de tracé de rayons.

Dans tout système de radiocommunication, les caractéristiques du canal de propagation sont essentielles : elles dépendent de la fréquence, des environnements, de l'emplacement des stations, etc ... L'optimisation des traitements effectués aussi bien à l'émission qu'à la réception pour compenser les effets perturbateurs de la liaison (atténuation, dispersion temporelle, fluctuations, ...) est nécessaire pour que le système puisse résister aux bruits tout en maintenant une capacité maximum. Il en résulte une approche statistique fondées sur des mesures du canal de propagation : puisque les effets des multitrajets sont linéaires, le canal est représenté comme un filtre linéaire dont la réponse impulsionnelle $h(t, \tau)$ varie avec le temps ; en outre, comme il est difficile de représenter un grand nombre de multitrajets, $h(t, \tau)$ est assimilée à un processus stochastique.

Grâce à l'expérience acquise en ces deux domaines et au perfectionnement des modèles, on peut imaginer une convergence des deux approches vers un modèle de base commun servant à déterminer les conditions précises du fonctionnement harmonieux des radars à large bande avec les autres services. Ce serait un premier pas nécessaire d'un très long chemin vers le partage de bandes.

REFERENCES BIBLIOGRAPHIQUES

[1] Anwendungsmöglichkeiten von Niederfrequenzradar-anlagen
Jean ISNARD
in ELOKA-Symposium pp 25-1 / 25-7
Bundesakademie für Wehrverwaltung und Wehrtechnik
Mannheim, 7-9 Oktober 1994

[2] Introduction to ultra-wideband radar systems
edited by James D. TAYLOR
CRC Press, Inc. 1995

[3] OSD / DARPA, Ultra-Wideband Radar Review Panel, Assessment of Ultra-Wideband (UWB) Technology
DARPA, Arligton, VA, 1990
The executive of summary this OSD/DARPA sponsored study can be found in the November, 1990 issue of the IEEE AES magazine on pp 45 to49

[4] Dimensionalized wavelet transform with application to radar imaging
J. BERTRAND, P. BERTRAND, JP. OVARLEZ
in Proc. IEEE - ICASSP, Toronto Canada
May 1991, pp 2909-2912

[5] Unambiguous range-Doppler processing method and system
Paul Van ETTEN, Michael WICKS
United States Patent Number 5, 657, 022
August 12, 1997

[6] Properties of Electromagnetic beams generated by Ultra-wide Bandwidth pulse-driven arrays
R. ZIOLKOWSKI
in IEEE Trans on A & P Vol. 40, N° 8 August, 1992
pp 888 - 905

[7] Appendix S3 (Section II) of the Radio Regulations (RR) of the ITU

[8] Les applications du spectre radioélectrique
Paul PIGNAL, François RANCY, Eric FOURNIER, Martine LIMODIN, Christophe de MAILLAR
in Actes Forum du Conseil Scientifique de FRANCE TELECOM Memento N° 6 (26/10/95) pp 9-23

[9] An ultra wide-band coherent radar system optimizing the radio interference discrimination capability
Hans HELLSTEN
Brevet International Publication Number WO 97/36188
2 October 1997

[10] Radar signal optimisation for VHF-experimental radar LARISSA
H. KUSCHEL
Colloque International sur le radar
Radar 94, Paris May 3-6, 1994 pp 100-104

PAPER TITLE : Paper 16 - Conditions d'insertion des radars à large bande

AUTHOR : J. Isnard

NAME : D. Jaeger

QUESTION :

Is there any tendency for the UWB radar with respect to resistance against jamming? Will it become better or worse than conventional radars?

ANSWER :

The UWB radar will have to be better in terms of resistance to jamming, the tendency is to associate more and more radar and ELINT equipment from the front-end to the signal processing. For example dynamic range of 80 dB will be obtained in the near future. Obviously the front-end (including the aerials) is the essential part of this evolution in radar technology.

NAME : I. White

QUESTION :

What are the peak and mean powers of this radar?

What is the cost of band spectral leakage?

ANSWER :

L'ordre de grandeur de puissance générée est d'une centaine de Watts.

On peut atteindre des concentrations spectrales de l'ordre de 40 à 50 dB.

PAPER TITLE : **Paper 16 - Conditions d'insertion des radars à large bande**

AUTHOR : J. Isnard

NAME : M. Elliott

QUESTION :

I think most people here will agree that the future for military equipment is wider-band bands and wide timing ranges.

You indicated that only the 2.7-3.4 GHz and 8.5-10.5 GHz bands would provide the spectrum needed for these very wide bandwidth radars. When do you think we could expect to see these radars introduced?

ANSWER :

As far as X band (8-12 GHz) is concerned we will probably have surely a mode in the next ten years for airborne radars because the target discrimination/ identifications is so important.

In N band (2.7/3.4 GHz) I am not so sure but it seems to me that for shipborne applications it should be equally important to improve the capacity of identification.

AIRBORNE METRIC FREQUENCY SURVEILLANCE RADAR (UHF - VHF)

Ph Lacomme - B. Carrara

THOMSON-CSF / Radars & Contre Mesures

La Clef de Saint Pierre
1, boulevard Jean Moulin
78852 Elancourt Cedex

This paper presents a new concept of airborne surveillance system using a low frequency band (metric wave length) aimed at :

- detecting and tracking air targets with low Radar Cross Section (RCS) such as stealth aircraft or missiles,
- detecting and localising moving and non-moving ground targets, possibly hidden under foliage.

1. AIRBORNE TARGETS DETECTION

Current systems have been designed to detect air targets whose RCS is relatively high (from some m^2 for classical aircrafts to some 0.1 m^2 for missiles). In front of low RCS targets, these systems are limited by :

- a loss of the power budget : a reduction of the RCS by a factor of 100 reduces the range by a factor 3
- an increase of the clutter to signal ratio. The radar improvement (and the subclutter visibility) and the antenna pattern quality become insufficient to detect the low altitude targets even at a shorter range. The global time to react is too short for the air defence efficiency.

Among the antistealth techniques, the use of low frequencies (< 500 MHz) is the most efficient. Absorbing material (RAM) efficiency is limited to short waves because their thickness is linked to the wavelength and active methods of nulling which are very sensitive to frequency and direction are not credible for airborne applications between 0.1 and 1 GHz as shown on figure 1.

With VHF waves the RCS of a stealth aircraft or a missile is higher than 0.1 m^2. Moreover for these wavelengths the ground backscattering coefficient is lower and the signal to clutter is higher.

2. GROUND TARGETS DETECTION

The propagation of high frequencies (> 500 MHz) suffers from high attenuation by foliage which avoids detection of fixed or mobile ground targets hidden by the vegetation.

The foliage penetration properties of low frequencies have been demonstrated by various experiments among which SACHEM experiments (THOMSON-CSF Applications Radar TCAR). As an example reference [2] presents measurements made with the radar CARABAS of the National Swedish Defence Research Establishment and with the International SRI Ultra-Wideband SAR. These measures focus on two types vegetations, the Maine forest and a rain forest of Panama. They show that the two way attenuation due to the foliage is acceptable (some decibel) in the 200 MHz- 400 MHz band, but becomes very strong (> 15 dB) above 1 GHz.

3. ISSUES FOR METRIC FREQUENCIES RADARS

However the use of such frequencies raise the issues of :

- international regulations (CCIR),
- airborne platform installation.

Radar designers have abandoned metric waves a long time ago for higher frequencies (>1 GHz), but distribution tables show that potential free bands could exist due to communication technical evolution.
These bands, some of which are reserved to Military Forces, are very sought after and must be strongly claimed for radar applications.

Paper presented at the RTO IST Symposium on "Frequency Assignment, Sharing and Conservation in Systems (Aerospace)", held in Aalborg, Denmark, 5-7 October 1998, and published in RTO MP-13.

The high density spectrum in these bands raises the hard interference issue (co-existence with other systems present on the platform or extern to it). Efficient filtering techniques will be needed on transmission (limitation of the transmitted spectrum) and on receive (spatial or time adaptive rejection of polluting signals).

In airborne radar, the combined effect of lower radiation performance (low quality antenna pattern) for UHF/VHF bands compared to higher band (S,X...) and the carrier motion , spread the ground clutter spectrum in whole the useful target Doppler domain.

In order to overcome this problem new techniques for compensation of the platform motion (Immobilisation of the antenna phase centre: DPCA, adaptive cancellation of the clutter..) will be necessary.
They enable to limit the ground clutter spectral spreading, facilitate the ground echo cancellation and improve the detectability of low velocity targets.

STAP processing is the object of an intense activity in laboratories and publications are numerous. Preliminary studies undertaken have provided promising results. Nevertheless validation works are required with an airborne demonstrator.

SAR techniques (Synthetic Aperture Radar) allowing to reach the transverse resolution needed for mapping of ground and fixed targets are perfectly mastered. On the other hand the bandwidth required to obtain the radial resolution (15 MHz) raises the problem of the utilisation of the spectrum, underlined hereafter. Techniques of "step frequency" will be used to restore this band (or a part of it) from disconnected subbands and will be complementary to coherent tomography.

One of the major risks of a such concept is the practical possibility of using the few available radar windows in this overloaded part of the spectrum. It would be nevertheless so prejudicial that designers and users of radars abandon, without combat, these bands of frequency whose radar properties are so particular.

4. CONCLUSION

The operational interest presented by such systems justifies that organisms implied in the frequencies allocation restore to the radar the access to VHF and UHF bands too rapidly abandoned.

A LINEAR PROGRAMMING APPROACH TO RADIO CHANNEL ASSIGNMENT IN HEAVILY LOADED, EVOLVING NETWORKS

Robert A. Leese
Mathematical Institute
University of Oxford
Oxford OX1 3LB, UK

Summary

This paper investigates a possible approach to the radio channel assignment problem, based on linear programming relaxation, together with column generation. The method is tested on a set of benchmarks that are expected to be challenging, and most cases are handled well. Those that are not suggest possible improvements. The method becomes more attractive when there are multiple channel demands at each transmitter site. Attention is restricted to minimum span problems, with interference controlled by a constraint matrix, but similar approaches are possible for more general formulations.

This research was funded by the UK Radiocommunications Agency, under a contract with St. Catherine's College, Oxford.

1 Introduction

Radio channel assignment, when treated in any general setting, is a hard combinatorial problem, with little prospect of guaranteeing to calculate optimal assignments in reasonable time. Instead, one aims for methods that are fast enough for real applications and that produce assignments sufficiently close to optimal, as measured by a suitable performance metric (one common measure being the number of channels used).

This paper discusses some refinements to an algorithm that in an earlier form has produced promising results for a variety of standard benchmarks. The method is based on a formulation in terms of linear programming, for which there are fast and reliable computational algorithms (the current implementation uses the well-established simplex method). The main new ingredient here is the idea of 'column generation', which means roughly that the complete problem formulation need not be compiled before the solution can begin. Instead, the formulation is augmented as the computation proceeds. In this way, redundant parts of the full formulation never need to be considered. Column generation has been used previously for a variety of combinatorial problems, notably for the graph-colouring problem [1], and also recently in a different approach to the channel assignment problem [2]. We shall refer to the method presented here as the CACG (Channel Assignment by Column Generation) procedure.

The CACG procedure may be applied to the range of problems discussed in the tutorial paper by Smith *et al.* [3]. Initial results for well-established benchmark problems have been reported in [4]. In this paper we shall look instead at some problems based on regular geometries, since these appear to be among the most difficult for general algorithms to deal with. The problem domain is the region of the plane shown in figure 1, covered by a mesh of 37 regular hexagonal cells, each with a transmitter site at its centre. Initially, each site is to be assigned a single channel, although we will also describe results for multiple coverage problems, in which each site receives several channels. The transmitters are identical and omnidirectional, signal propagation is assumed to be isotropic and independent of frequency, and no terrain effects are considered. These assumptions give rise to a high degree of symmetry in the problem, which algorithms must 'discover' for themselves if they are to produce good assignments. It is a challenge to design algorithms that can deal equally well with both highly symmetric problems and more irregular ones. Surprisingly, perhaps, it seems that symmetric problems are harder to handle.

Quality of service is guaranteed by controlling interference coming from signals transmitted on channels that are close in frequency to the wanted signal. In engineering specifications, these guarantees are usually provided by protection ratios. Mathematically, it is more common to use a constraint matrix, which is part of the following general framework.

Suppose that the set of transmitter sites is $T = \{t_i : i = 1, 2, \ldots n\}$, and that transmitter site t_i should be assigned m_i channels. In keeping with standard notation, the available channels will be denoted by

Paper presented at the RTO IST Symposium on "Frequency Assignment, Sharing and Conservation in Systems (Aerospace)", held in Aalborg, Denmark, 5-7 October 1998, and published in RTO MP-13.

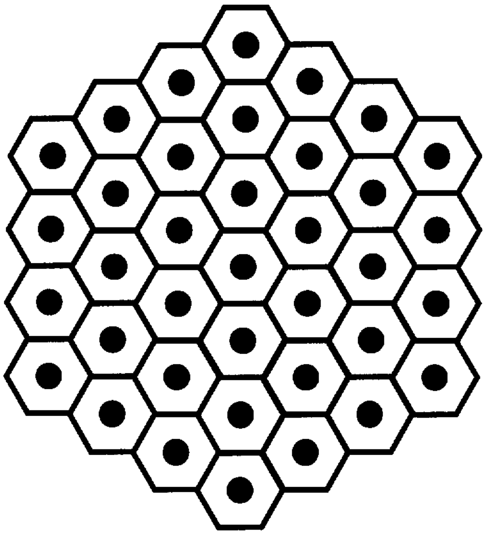


Figure 1: An arrangement of 37 regular cells, each with a transmitter site at the centre. The distance between nearest sites is one unit.

$0, 1, 2, \ldots$. The mapping between channels and sites determines an assignment.

Definition 1 *An* assignment *A on the set T, is a (finite) subset of $T \times \{0, 1, 2, \ldots\}$. Its* span *is $max\{f : (t, f) \in A\}$.*

In other words, an assignment is a collection of ordered pairs (t, f), each of which indicates that channel f is assigned to site t. The number of channels $m_A(t)$ provided at t by the assignment A may be written

$$m_A(t) = |\{f : (t, f) \in A\}| . \qquad (1)$$

Definition 2 *The assignment A is* consistent *with the constraint matrix $C = (c_{ij})$ if $|f_i - f_j| \geq c_{ij}$ for all distinct $(t_i, f_i), (t_j, f_j)$ in A.*

Definition 3 *The* linear span, *denoted χ_{lin}, is the smallest span S of any assignment on the transmitter sites T that is consistent with the constraint matrix C and has $m_A(t) \geq m_i$ for all $t_i \in T$.*

The reason for attaching the adjective 'linear' and subscript 'lin' in definition 3 is that we shall also consider the *cyclic span*, which is defined slightly differently.

Definition 4 *The assignment A (of span $S - 1$) is* cyclically consistent *with the constraint matrix $C = (c_{ij})$ if $|f_i - f_j|_S \geq c_{ij}$ for all distinct $(t_i, f_i), (t_j, f_j)$ in A, where*

$$|f_i - f_j|_S = \min(|f_i - f_j|, S - |f_i - f_j|). \qquad (2)$$

Definition 5 *The* cyclic span, *denoted χ_{cyc}, is the smallest value of S for which there exists an assignment of span $S - 1$ on the transmitter sites T that is cyclically consistent with the constraint matrix C and has $m_A(t) \geq m_i$ for all $t_i \in T$.*

Note that, in definitions 4 and 5, the highest channel is $S - 1$, not S as before; equation (2) effectively places the channels around a circle, with 0 and $S - 1$ adjacent. The use of $S - 1$ as the highest channel is a matter of convention. It was used in [5], where only cyclic spans were considered. Moreover, it is very natural when considering problems of multiple coverage [6]: one way of providing several channels at each site is to lay down copies of an assignment with cyclic span S, beginning with channels $0, S, 2S, \ldots$. The cyclic span is then the number of extra channels used to provide an additional channel at each site. (Note that in [6], $S - 1$ was used as the highest channel in defining the linear span, but here we have chosen S as the highest channel, to prevent conflict with the earlier tutorial. As a consequence, in this paper the cyclic span is always at least one more than the linear span.)

In regular transmitter geometries such as figure 1, the constraint matrix C may be replaced by a set of constraint distances $d_0, d_1, d_2, \ldots$. Each d_i is the smallest allowed spatial separation of transmitter sites that may be assigned channels i apart in the spectrum. Hence d_0 constrains channel re-use, d_1 the use of first adjacent channels, and so on. Here we shall consider constraints imposed on channel separations 0, 1 and 2 through the specification of three distances (d_0, d_1, d_2). On a uniform mesh, the d_i may, without loss of generality, be restricted to values that are possible distances between pairs of sites. For hexagonal cells (and taking the distance between nearest cell centres to be one unit), each d_i^2 must be a rhombic number (see, for example, [7]), *i.e.* a number that may be written in the form $p^2 + pq + q^2$ for some integers p and q. The first few rhombic numbers are $1, 3, 4, 7, 9, 12, 13, 16, 19, 21$. To keep the constraint parameters integer-valued, it is convenient to use the squared distances (d_0^2, d_1^2, d_2^2), with the linear and cyclic spans denoted by $\chi_{\text{lin}}(d_0^2, d_1^2, d_2^2)$ and $\chi_{\text{cyc}}(d_0^2, d_1^2, d_2^2)$, respectively. For any set of such constraints,

$$\begin{aligned} &\chi_{\text{lin}}(d_0^2, d_1^2, d_2^2) \\ &< \chi_{\text{cyc}}(d_0^2, d_1^2, d_2^2) \\ &\leq \chi_{\text{lin}}(d_0^2, d_1^2, d_2^2) + 3 . \end{aligned} \qquad (3)$$

We shall look here at the twenty problems corresponding to (d_0^2, d_1^2, d_2^2) being one of the triples of rhombic numbers satisfying $1 \leq d_2^2 < d_1^2 < d_0^2 \leq 12$.

These constraints have previously been considered in the limit of an infinite two-dimensional mesh of hexgonal cells [6]. The values of $\chi_{\text{lin}}(d_0^2, d_1^2, d_2^2)$ and

$\chi_{\text{cyc}}(d_0^2, d_1^2, d_2^2)$ are known in the case when a single channel is required at each site. Their calculation is based on exhaustive searches using small sets of cells, to provide lower bounds on the spans. By adding more cells to the search domain (for a given set of constraints), the lower bound initially increases, but eventually a pattern of channels emerges that can be extended to an assignment on the infinite mesh, without needing to introduce any extra channels. Consequently, the best lower bound is then confirmed to be at the same time an upper bound, and hence equal to the span itself.

It should be clear that these methods do not constitute a satisfactory algorithm. First, the searches that establish lower bounds are very time-consuming; secondly, human intervention is used to spot the patterns that establish upper bounds. However, they do give a set of useful benchmarks. For each of the twenty problems under consideration here, the work described in [6] showed that the linear and cyclic spans on the region shown in figure 1 are equal to the corresponding spans on the infinite mesh. The values given in [6] are reprinted here in the third and fourth columns of table 1 (with the linear span reduced by one, owing to the slight change in definition mentioned above).

The remainder of the paper is arranged as follows. Section 2 describes the components in the CACG procedure, and in particular the way in which column generation works for this problem. Section 3 presents results when a single channel is needed at each site (for reasons that will be discussed, these are expected to be the cases in which CACG performs worst), and compares them to the optimal results in table 1. Section 4 discusses the operation of CACG when the demand at each site is increased to five channels, and section 5 presents conclusions.

2 The CACG procedure

2.1 Overview

The CACG procedure is a combination of several standard techniques from linear optimization. The following paragraphs give a brief description, but the reader might also wish to acquire further details from a modern text in the field, such as [8].

Many problems in combinatorial optimization may be written as integer programs, *i.e.* as the maximization or minimization of a linear objective subject to linear constraints and with the restriction that all variables be integer valued. A common first step is to 'relax' to a linear program by removing the restriction of integer values. CACG follows this route. The linear program is generally much easier to solve, but produces a fractional solution, which is therefore not feasible for the original problem. Two ways of reverting to integer values are to apply a branch-and-bound procedure or simply to round the fractional values. The former finds an optimal solution to the original problem, but can be time-consuming; the latter is fast, but generally suboptimal and may cause some constraints to be violated. In CACG, rounding is used, since we have one eye on speed, and the approximations introduced are not expected to be severe compared with those from other sources.

Column generation is applied to the linear programming relaxation of the original problem. The general strategy involves a formulation in terms of small 'building blocks'. For example, in the graph-colouring problem, the natural building blocks are independent sets. In CACG, the building blocks are generalizations of independent sets, known as spectral blocks. An 'auxiliary problem' is used to construct the building blocks. As many as possible of the constraints in the original problem are moved to the auxiliary problem. Each building block corresponds to a variable in the linear program, and so to a column in the associated matrix of linear constraints. The blocks produced by the auxiliary problem add new columns to the original problem, hence the term 'column generation'.

The method of column generation is particularly attractive when there are potentially a very large number of variables in the original problem. For example, the number of independent sets in a graph is typically exponential in the number of nodes, and similarly with the number of spectral blocks in a channel assignment problem. The original and auxiliary problems are considered alternately, with new blocks (columns) being introduced by the latter based on the current state of the former. In this way, irrelevant variables are never introduced. If the intention is not necessarily to reach an optimal solution then a suitable termination criterion can be applied.

Let us now see how channel assignment fits into this framework. In this paper, attention is confined to minimum span problems (see [3]), but fixed spectrum problems are amenable to a similar treatment.

Definition 6 *A* spectral block B of size λ *is an assignment of span* $\lambda - 1$ *that is cyclically consistent with the constraint matrix* C*, and has the additional properties* $m_B(t_i) \leq 1$ *for all* $t_i \in T$ *and* $(t_i, f_i) \in B \Rightarrow c_{ii} \leq \lambda$.

Think of a spectral block B as making available a single channel to each site in some subset of T, subject to the cyclic interference constraints. Recalling the relevance of the cyclic span to problems of multiple coverage, we see that a block B may be laid

Instance	(d_0^2, d_1^2, d_2^2)	χ_{lin}	χ_{cyc}	S^*	$\hat{S}$	N_u
1	$(4,3,1)$	6	8	7	8	—
2	$(7,3,1)$	8	9	10	9	—
3	$(7,4,1)$	12	13	14	12	—
4	$(7,4,3)$	13	14	14	12	—
5	$(9,3,1)$	10	11	10	9	—
6	$(9,4,1)$	12	13	14	12	—
7	$(9,4,3)$	13	14	14	12	—
8	$(9,7,1)$	16	18	18	18†	14
9	$(9,7,3)$	18	20	18	18†	15
10	$(9,7,4)$	20	21	26	21	—
11	$(12,3,1)$	11	12	11	12	—
12	$(12,4,1)$	13	14	14	12	—
13	$(12,4,3)$	13	14	14	12	—
14	$(12,7,1)$	17	18	19	20.0194†	15
15	$(12,7,3)$	19	20	22	20.8462†	7
16	$(12,7,4)$	20	21	26	21	—
17	$(12,9,1)$	22	24	23	24	—
18	$(12,9,3)$	22	24	22	24.5106†	12
19	$(12,9,4)$	25	28	26	24†	11
20	$(12,9,7)$	29	30	33	27	—

Table 1: The twenty benchmark problems with unit demand. The third and fourth columns give the linear and cyclic spans; S^* is the value output by CACG, and should be compared against χ_{lin}; $\hat{S}$ is the linear programming objective in CACG (see section 2.3), with fractional solutions marked by a dagger; in the case of fractional solutions, N_u is the number of unsatisfied demands handled by postprocessing.

down in the frequency spectrum several times, starting at channels $0, \lambda, 2\lambda, \ldots$, thereby providing several channels at each site.

A complete assignment is made up from a sequence of blocks, as shown in figure 2, with each block B_b (of size λ_b) repeated μ_b times before a new block is used. To satisfy all interference constraints, it is possible that 'guard intervals', consisting of one or more unused channels, need to be inserted between the last repetition of one block and the first repetition of the next. In the current implementation of CACG, guard intervals are added in a post-processing phase, and so for the moment we shall ignore them.

2.2 Linear programming formulation

Given the (exponentially large) set $\mathcal{B} = \{B_b : b \in I\}$ of possible spectral blocks, indexed by a set I, the minimum span problem may be written as

$$\begin{aligned} \text{(IP): minimize} \quad & \textstyle\sum_{b \in I} \lambda_b \mu_b \\ \text{subject to} \quad & \textstyle\sum_{b \in I} A_{bi} \mu_b \geq m_i \ (t_i \in T) \\ \text{and} \quad & \mu_b \text{ a non-negative integer } (b \in I). \end{aligned}$$

Here, A_{bi} is equal to 1 if $(t_i, f_i) \in B_b$ for some f_i and zero otherwise.

The next stage is to write a linear programming relaxation, by replacing the integer variables μ_b with real variables $\hat{\mu}_b$:

$$\begin{aligned} \text{(LP): minimize} \quad & \textstyle\sum_{b \in I} \lambda_b \hat{\mu}_b \\ \text{subject to} \quad & \textstyle\sum_{b \in I} A_{bi} \hat{\mu}_b \geq m_i \ (t_i \in T) \\ \text{and} \quad & \hat{\mu}_b \geq 0 \ (b \in I). \end{aligned}$$

It is useful to consider the linear programming dual of (LP), namely

$$\begin{aligned} \text{(D): maximize} \quad & \textstyle\sum_i m_i y_i \\ \text{subject to} \quad & \textstyle\sum_i A_{bi} y_i \leq \lambda_b \ (b \in I) \\ \text{and} \quad & y_i \geq 0 \ (t_i \in T). \end{aligned}$$

The dual problem has the same optimal value as (LP), and provides a better setting in which to discuss the column generation process. Suppose that we have a (small) set of blocks $\mathcal{B}_j \subset \mathcal{B}$, and consider the restricted problem

$$\begin{aligned} (\text{D}_j)\text{: maximize} \quad & \textstyle\sum_i m_i y_i \\ \text{subject to} \quad & \textstyle\sum_i A_{bi} y_i \leq \lambda_b \ (B_b \in \mathcal{B}_j) \\ \text{and} \quad & y_i \geq 0 \ (t_i \in T). \end{aligned}$$

Column generation is an iterative process in which (D_j) is solved for $j = 0, 1, 2, \ldots$, with one or more

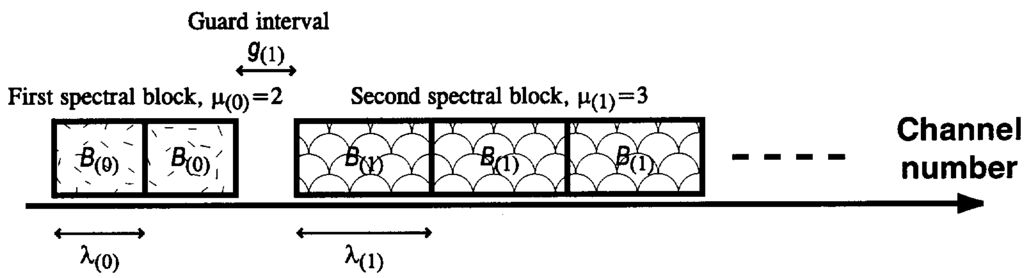


Figure 2: The general way in which an assignment is made up of spectral blocks, as discussed in section 2.1. The subscripts in parentheses relate to block ordering (see section 2.3.2).

new blocks added to $\mathcal{B}_j$ at each iteration to form $\mathcal{B}_{j+1}$. The best blocks to add are those that violate the corresponding constraint in (D). They are found by considering the following auxiliary problem. Suppose that at the current iteration the values y_i^* are an optimal solution to (D_j). Choose a size λ' for the new block B' and attempt to

(AUX): maximize $\sum_i a_i y_i^*$

subject to

(a_i) being the incidence vector for B'.

Here the variables are the a_i, and by saying that they make up an incidence vector for B' we mean that a_i is 1 if $(t_i, f) \in B'$ for some f and zero otherwise. In other words, one looks for blocks of large overall 'weight' where the weight associated with site t_i is y_i^*, as determined by the most recent solution of (D_j). In practice, one might wish to run (AUX) several times for different values of λ'. Moreover, it is not strictly necessary to find optimum solutions to (AUX). Any blocks B' with overall weight greater than λ' correspond to constraints in (D) that are violated by the current solution to (D_j). Adding them to $\mathcal{B}_j$ provides a 'cutting plane' [8] in the search for the common optimum value of (LP) and (D).

In the current implementation, (AUX) is approached by putting the sites in order of decreasing weight y_i^*, and then performing a depth-first search on a binary tree. A node at level j in the tree corresponds to the choice of whether or not to attempt to include in B' the site of jth highest weight. To be explicit, it is useful to have the following definition (where the associated constraint matrix is understood).

Definition 7 *Two (cyclically) consistent assignments, A_1 and A_2, of spans S_1 and S_2 respectively, are* (cyclically) compatible *if $A_1 \cup A_2$ is a (cyclically) consistent assignment of span* $\max(S_1, S_2)$.

When attempting to add a site t to B', we look to replace B' with $B' \cup \{(t,f)\}$ where f is the smallest value in $\{0, 1, \ldots, \lambda' - 1\}$ such that $\{(t,f)\}$ is cyclically compatible with B'. If there is no possible f then the search tree is pruned at that point and a backtracking step is made to the parent node. Even so, when there are many non-zero values of y_i^*, the complete search may be too lengthy. Bearing in mind that we need to find blocks of high total weight, but not necessarily the highest possible, a time cutoff is imposed, at which the search is aborted and the block of highest weight found so far added to $\mathcal{B}_j$.

Before describing the post-processing phase, we mention breifly how the initial set $\mathcal{B}_0$ is found. Here, the main concern is that every $t \in T$ is served by some block in $\mathcal{B}_0$, so ensuring that (D_0) has a finite optimum value. This is easily accomplished by sequentially constructing blocks as follows.

$\mathcal{B}_0 \leftarrow \phi$
$V \leftarrow T$
$\lambda \leftarrow \max\{c_{ii} : t_i \in T\}$
while $V \neq \phi$
  $B \leftarrow \phi$
  while $\exists (t,f) \in V \times \{0, 1, \ldots \lambda - 1\}$, cyclically compatible with B
    $B \leftarrow B \cup \{(t,f)\}$
    $V \leftarrow V \setminus \{t\}$
  $\mathcal{B}_0 \leftarrow \mathcal{B}_0 \cup B$
end

2.3 Postprocessing

Suppose that the column generation procedure results in a value $\hat{S}$ for the linear program (LP), and corresponding variables $\hat{\mu}_b$. The simplex method will always output a set of $\hat{\mu}_b$ containing no more than $|T|$ nonzero values. In general some of them will be fractional, so they are not a feasible solution for (IP). Even when all the $\hat{\mu}_b$ are integer-valued, the corresponding blocks have yet to be assembled into an overall assignment. To address these issues, up to four postprocessing stages are employed:

1. Conversion to an integer solution

2. Block ordering and insertion of guard intervals

3. First provision for unsatisfied demand

4. Final provision for unsatisfied demand.

If all the $\hat{\mu}_b$ are integer valued, then only the second of the four stages is needed. Each stage is briefly described below, but first it is useful to have the following notion of 'adding' two assignments.

Definition 8 *If assignments A_1 and A_2 have spans S_1 and S_2, respectively, then their* sum $A_1 + A_2$ *is defined to be* $A_1 \cup \{(t, f + S_1 + 1) : (t, f) \in A_2\}$ *and has span* $S_1 + S_2 + 1$.

Note that this addition is not commutative. Also, $A_1 + A_2$ is not necessarily consistent, even if A_1 and A_2 are both consistent. We can now move on to a discussion of postprocessing.

2.3.1 Integer solutions

The fastest way of obtaining integer solutions is to round any fractional $\hat{\mu}_b$. A key consideration with rounding schemes is to deal with any constraints that might be violated as a result. Here, we could guarantee to violate none of the constraints in (IP) by rounding every $\hat{\mu}_b$ up, but the resulting span is then generally poor. Instead, we adopt randomized rounding [9], giving integer values μ_b^*: if $\hat{\mu}_b$ is integer-valued, then set $\mu_b^* = \hat{\mu}_b$, otherwise make the random choice

$$\mu_b^* = \begin{cases} \lceil \hat{\mu}_b \rceil & \text{with probability } \hat{\mu}_b - \lfloor \hat{\mu}_b \rfloor \\ \lfloor \hat{\mu}_b \rfloor & \text{with probability } \lceil \hat{\mu}_b \rceil - \hat{\mu}_b. \end{cases} \quad (4)$$

Postprocessing is sufficiently fast that it may be executed several times with independent choices for the randomized rounding, and then the lowest of the resulting spans taken as the final output.

2.3.2 Block ordering

The rounding procedure leaves us with a set of blocks $\mathcal{B}^*$ that will be used to make up the final assignment, together with corresponding multiplicities μ_b^*, but does not dictate the order in which the blocks are to appear in the frequency spectrum (see figure 2). Suppose the elements of $\mathcal{B}^*$ are ordered $(B_{(0)}, B_{(1)}, \ldots, B_{(r)})$, with respective sizes $(\lambda_{(0)}, \lambda_{(1)}, \ldots, \lambda_{(r)})$ and multiplicities $(\mu_{(0)}, \mu_{(1)}, \ldots, \mu_{(r)})$, where $r = |\mathcal{B}^*| - 1 < |T|$. Then they may be combined iteratively into a single assignment, of the form shown in figure 2, by setting

$$\begin{aligned} A_{(0)} &= \{(t, f + j\lambda_{(0)} - \delta) : (t, f) \in B_{(0)}, \\ &\qquad 0 \le j < \mu_{(0)}\} \\ A_{(i)} &= A_{(i-1)} + \{(t, f + j\lambda_{(i)} + g_{(i)}) : (t, f) \in B_{(i)}, \\ &\qquad 0 \le j < \mu_{(i)}\},\ 1 \le i \le r. \end{aligned} \quad (5)$$

Here, $\delta = \min\{f : (t, f) \in B_{(0)}\}$, effectively removing any unused channels at the start of the assignment; $g_{(i)}$ is the guard interval inserted between the last repetition of $B_{(i-1)}$ and the first repetition of $B_{(i)}$, chosen as small as possible so that $A_{(i)}$ is consistent.

The block ordering procedure used in the current implementation is straightforward. First, $B_{(0)}$ is chosen to maximize δ, and then each subsequent $B_{(i)}$ is chosen to minimize $g_{(i)}$. Where there are several equally good candidates, one of them is chosen arbitrarily. Ideally, the $g_{(i)}$ account for only a small fraction of the final span. Since the number of guard intervals is no more than the number of sites, they will become less significant as the channel demands m_i (and hence also the block multiplicities μ_b^*) increase.

The only refinement concerns the internal structure of the block, *i.e.* the details of precisely which channel is assigned to which site. Note that the linear program looks only at the incidence vector of each block, not at the internal details, which are all dealt with by the auxiliary problem (AUX). However, it may be possible to reduce guard intervals by keeping the same incidence vectors but modifying the internal structure. Investigating all possible internal structures consistent with the interference constraints is probably not worth any small gain that might result, but we do allow blocks to be 'rotated' in the attempt to find small guard intervals. If a block B has size λ, then its rotation B^a through a steps is defined to be $\{(t, f + a \pmod{\lambda})) : (t, f) \in B\}$. Rotation preserves consistency with the interference constraints.

2.3.3 First provision for unsatisfied demand

Suppose we let A', with a linear span S', be the assignment $A_{(r)}$ given by the final iteration in the block ordering procedure. The earlier rounding of the $\hat{\mu}_b$ may lead to violated constraints in (IP), *i.e.* sites t_i for which $m_{A'}(t_i) < m_i$, in which case extra channels need to be found to satisfy the full demand. First, an attempt is made to do this without increasing the span. As long as there exist pairs (t_i, f_i) for which $m_{A'}(t_i) < m_i$, $f_i \le S'$ and $\{(t_i, f_i)\}$ is compatible with A' then we can replace A' with $A' \cup \{(t_i, f_i)\}$.

2.3.4 Final provision for unsatisfied demand

Suppose after the first provision for unsatisfied demand we have an assignment A'' (still with span S'). Any remaining unsatisfied demands are met by adding extra channels beyond S'. There are generally sufficiently few such demands that an exhaustive search is very fast. Explicitly, we find an assignment A_u (subscript u for 'unsatisfied') of span S_u such that $m_{A_u}(t_i) = m_i - m_{A''}(t_i)$ for all $t_i \in T$ and $A'' + A_u$ is consistent. This is the final output from the CACG procedure as currently implemented.

3 Results with unit demands

This section discusses the performance of CACG on problems that are expected to be among those less well suited to its application. They all require the assignment of a single channel at each of the 37 sites in figure 1, subject to one of the 20 different sets of interference constraints previously identified in table 1. The fifth and sixth columns in that table show the span S^* of the final assignment output by the CACG procedure, and also the (possibly fractional) value $\hat{S}$ found by the underlying linear program. The differences between the two reflect the effects of the various postprocessing stages described in section 2.3. A dagger against a value of $\hat{S}$ indicates that some of the corresponding $\hat{\mu}_b$ are fractional and hence that full postprocessing has been carried out. For these cases, the final column shown the number of unsatisfied demands handled in the postprocessing stage.

For all calculations, the initial set of blocks, $\mathcal{B}_0$, contained blocks of size 3 only. The auxiliary problem looked for blocks of sizes 3 and 4, and, for each size, the maximum time allowed for the search was 1 second (on a Pentium processor running at 150 MHz). The column generation procedure was terminated at the first iteration where no blocks were discovered with incidence vectors different to what had already been included. For fractional solutions, ten randomised roundings were performed and the lowest span recorded. The longest total execution time was approximately 5 minutes. The longer calculations occurred for the fractional cases, often with much time spent in the final stages of column generation to make small improvements in $\hat{S}$. Since these small improvements are probably 'rounded away' in postprocessing, there seems scope for a better termination criterion. Apart from the very fast instances, about 80% of time is spent in the column generation routine. Execution times do not seem to depend in any systematic way on the numerical values of the interference constraints.

The worst results are those for instances 10, 15, 16 and 20, where S^* differs from χ_{lin} by more than 2. The three worst cases have integer-valued solutions, suggesting that perhaps the block ordering procedure could be improved, so reducing the need for guard intervals. Alternatively, it may be that sometimes spectral blocks simply do not provide a good setting for the calculation: with unit demands, the cyclic interference constraints placed on each block are effectively unnecessary extra restrictions.

For the fractional cases, rounding tends to leave a high proportion of unsatisfied demands, again a legacy of unit demands. The symmetry in the problem leaves little scope for provision at the third stage of postprocessing (recall the third stage does not allow for increasing the span). Hence the final stage, by exhaustive search, plays a larger role than we would ideally like, although the problems here are small enough that the effects on execution time are insignificant. Very large regular problems with unit demand would incur much larger searches, and hence incur a severe time penalty; but these are problems for which a specially tailored algorithm would be more appropriate.

A word of warning: it is not a good idea to control the total time spent in column generation by looking solely at the rate of decrease in the objective value $\hat{S}$. This is because the linear programs (LP) and (D) are highly degenerate, *i.e.* there are many different solutions all with the same objective value, especially for problems with high degrees of symmetry such as we have here. Each generated column violates a constraint in (D) and hence renders the current solution infeasible, but the next solution found often has the same $\hat{S}$. The instances studied here often exhibit runs of several tens of solutions all with the same value. Effectively, we are waiting for column generation to break the symmetry in the original problem, before $\hat{S}$ can start to decrease.

In summary, results from CACG are generally good, with a few inferior results suggesting possible improvements. The detailed operation highlights several undesireable aspects that are expected for regular problems with unit demand.

4 Results with higher demands

We now consider the same 20 problems, but with the demand at each site increased to five channels, *i.e.* the total demand is for 185 channels. One advantage of CACG is that this does not increase the typical execution time. Moreover, it is expected that the postprocessing phase is now less important: guard intervals will be a smaller component in the final span, and the unsatisfied demands caused by rounding will be a smaller fraction of the total demand.

Table 2 shows the values of S^* and $\hat{S}$ using the same operational settings as in section 3. The co-site constraint (*i.e.* the minimum separation between channels assigned to the same site) was set to 3. In contrast to the case of unit demands, we do not know the optimal values for each problem, but we can compare with the upper bound $5\chi_{\rm cyc} - 1$. For integer-valued solutions in table 1, another upper bound comes from multiplying each of the $\hat{\mu}_b$ by 5 and keeping the same block ordering and guard intervals; this bound is shown as $U_{\rm CACG}$ in table 2, and matches the span calculated from scratch except for instance 11 (which is now fractional).

In 13 of the 20 instances, S^* is equal to or less than $5\chi_{\rm cyc} - 1$, and in 6 of those by more than 10%. However, instances 10, 15 and 16 again suggest that a better block ordering procedure could be found. Instance 11 also highlights a potential improvement, since we know from table 1 that an integer solution exists with $\hat{S} = 60$. It seems that the termination criterion is activating sooner than it should. Instance 14 may have the same problem, although here it is reassuring that the value of $\hat{S}$ in table 2 is less than 5 times that in table 1.

Some instances that had integer-valued solutions in table 1 are now fractional, and vice-versa. This is counterintuitive, since scaling all m_i in (D) and (D$_j$) by the same factor should not change the corresponding y_i^* at each stage of the column generation. Further experiments have even shown that when an instance is repeated with the same m_i, a previously fractional (integer-valued) outcome is sometimes replaced by an integer-valued (fractional) one! A possible explanation is that the arbitrary ordering in the column generation routine of sites with equal y_i^* leads to different spectral blocks being generated for different runs, even with the same input data. In other words, limitations on machine precision could provide an element of randomness that has macroscopic, though harmless consequences. There seems no compelling reason to prefer integer solutions over fractional ones, other than they require less postprocessing and are easier to scale up to higher demands.

5 Conclusions

This paper has presented the CACG approach to radio channel assignment, based on linear programming and column generation. It is designed to achieve good results in reasonable time without the need for fine-tuning of internal parameters. The method has been tested on problems that are expected to be particularly challenging, even though they are relatively small.

The general validity of a scheme such as CACG is a key issue. It has the potential to be a good method whenever there is an optimal assignment that may be constructed from spectral blocks. Suppose that, given a problem instance, CACG produces an output that is known to be bad. This may indicate that CACG is not good for such problems, or it may be caused by a suboptimal solution to the underlying integer program (IP), or it may be caused by inappropriate postprocessing. In the second and third cases, the algorithm can be refined, but otherwise the problem would be better suited to a different algorithm. The benchmarks used in this paper have revealed good performance for many instances, and also possible improvements, which should help to clarify the range of application of such methods.

An aspect of channel assignment that has not arisen here is the idea that the span might be determined only by the demands and constraints at a few sites. The remaining sites would then be less important, in the sense that they could be provided with additional channels without increasing the span. Suppose a site t_i is called 'saturated' if increasing m_i by one channel causes the span (or some approximation to it, such as the $\hat{S}$ calculated by CACG) to increase, and called 'unsaturated' otherwise. Saturated sites can be identified from the final values of the y_i^* using the general theory of linear programming. Realistic problems often have a relatively small number of such sites, owing to an inhomogeneous distribution of demands. In constrast, the problems considered here tend to have all sites saturated. Intuitively, this means that there are fewer optimal assignments than in problems of similar size but with fewer saturated sites. If there are fewer optimal assignments, it then seems reasonable that they would be harder to find using a general algorithm. (Moreover, problems with few optimal assignments are less likely to have ones that can be constructed from spectral blocks.) Hence one could assess the difficulty of problems by the relative numbers of saturated and unsaturated sites. Although the problems considered here are small, they are among the hardest small problems.

The following paragraphs identify specific questions, refinements and extensions.

- The termination criterion for column generation is a delicate matter, especially for highly symmetric problems. A balance must be struck between a reasonably short execution time and the need to find at least near-optimal solutions to (IP).
- The detailed internal structure of spectral blocks is not explicitly treated in the current implementation. The expectation is that if a spectrally inefficient block is generated then it will simply be ignored in the subsequent calculations. But even if there is no loss in quality of the final assignment, there is an increase in execu-

Instance	(d_0^2, d_1^2, d_2^2)	$5\chi_{\text{cyc}} - 1$	U_{CACG}	S^*	$\hat{S}$	N_u
1	(4,3,1)	39	39	39	40	—
2	(7,3,1)	44	46	46	45	—
3	(7,4,1)	64	62	62	60	—
4	(7,4,3)	69	62	62	60	—
5	(9,3,1)	54	46	46	45	—
6	(9,4,1)	64	62	62	60	—
7	(9,4,3)	69	62	62	60	—
8	(9,7,1)	89	—	91	90†	8
9	(9,7,3)	99	—	91	90†	7
10	(9,7,4)	104	110	110	105	—
11	(12,3,1)	59	59	63	63.1355†	21
12	(12,4,1)	69	62	62	60	—
13	(12,4,3)	69	62	62	60	—
14	(12,7,1)	89	—	102	98.1774†	15
15	(12,7,3)	99	102	102	100	—
16	(12,7,4)	104	110	110	105	—
17	(12,9,1)	119	119	119	120†	6
18	(12,9,3)	119	—	119	120	—
19	(12,9,4)	139	—	122	120†	10
20	(12,9,7)	149	141	141	135	—

Table 2: Results from the CACG procedure for 20 benchmark problems based on the transmitter site layout in figure 1, when 5 channels are assigned at each site. The third and fourth columns are upper bounds on the corresponding linear span; S^*, $\hat{S}$ and N_u have the same meanings as in table 1.

tion time. There may be scope for a random component in the column generation, instead of sticking rigidly with a site ordering based on the current values of the y_i^*.

- It is not clear to what extent the choice of spectral block sizes affects performance. It is easier to generate optimal small blocks than optimal large ones, but the cost may come later in the form of guard intervals. It may be difficult to do well with sets of transmitter sites that show a wide range of different co-site constraints.

- The interference constraints considered here are examples of *binary* constraints, in the sense that each one puts a restriction on the channels allowed at a single *pair* of transmitter sites. It may be [10] that more general constraints should be used, to model more closely the fundamental requirement that the signal-to-interference ratio at each possible receiver location should be sufficiently high. The CACG procedure could be extended in this direction, by modifying the definition of a consistent assignment. The notion of spectral blocks and their associated linear programming formulation would remain.

- The title of the paper mentions 'evolving networks', by which is meant the possibility of the demands changing over time. Like many algorithms, CACG can easily use current solutions as a starting point for the calculation of new solutions with new demands. However, an advantage of CACG is that the tracking of slack variables, as mentioned above, means that if a new demand is made at a previously unsaturated site then there is no calculation needed at all. In terms of the underlying linear program, a calculation is needed only when the m_i change sufficiently that the optimum value in (D_j) moves from one vertex of the feasible region to another.

References

[1] A.Mehrotra and M.A.Trick, *A column generation approach for graph coloring*, INFORMS Journal on Computing, Vol. 8, No. 4 (1996).

[2] B.Jaumard and T.Vovor, *A column generation approach for the exact solution of channel assignment problems*, Research report, Ecole Polytechnique de Montréal (1998).

[3] D.H.Smith *et al.*, *Frequency assignment: methods and algorithms*, these proceedings.

[4] R.A.Leese, *Structural features in the construction of good channel assignments*, Proceedings of the 14th International Wroclaw Symposium on Electromagnetic Compatibility, June 23–25, 1998, pp. 710–711.

[5] J.van den Heuvel *et al.*, *Graph labelling and radio channel assignment*, to appear in Journal of Graph Theory.

[6] R.A.Leese, *A fresh look at channel assignment in uniform networks*, Proceedings of the 12th International Zurich Symposium on Electromagnetic Compatibility, February 18–20, 1997, pp. 127–130.

[7] V.H.MacDonald, *The cellular concept*, Bell System Technical Journal, Vol. 58, No. 1, pp. 15–41 (1979).

[8] D.Bertsimas and J.N.Tsitsiklis, *Introduction to linear optimization*, Athena Scientific, Belmont, Massachusetts (1997).

[9] D.Bertsimas and R.Vohra, *Rounding algorithms for covering problems*, Mathematical Programming, Vol. 80, pp. 63–89 (1998).

[10] J.Bater, *Are there effective binary frequency separation constraints for frequency assignment coverage problems?*, these proceedings.

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's References	3. Further Reference	4. Security Classification of Document
	RTO MP-13 AC/323(IST)TP/1	ISBN 92-837-0003-1	UNCLASSIFIED/ UNLIMITED

5. Originator Research and Technology Organization
North Atlantic Treaty Organization
BP 25, 7 rue Ancelle, F-92201 Neuilly-sur-Seine Cedex, France

6. Title

Frequency Assignment, Sharing and Conservation in Systems (Aerospace)

7. Presented at/sponsored by

The Information Systems Technology Symposium held in Aalborg, Denmark, 5-7 October 1998.

8. Author(s)/Editor(s)	9. Date
Multiple	January 1999

10. Author's/Editor's Address	11. Pages
Multiple	224

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

Information transfer	Radio frequencies
Frequency allocations	High frequencies
Frequency control	Electromagnetic compatibility
NATO forces	Optimization
Military communication	Neural nets
Communication and radio systems	Spread spectrum
Spectrum signatures	Frequency modulation
Systems management	Communications networks
Algorithms	Multiple access

14. Abstract

This volume contains the Technical Evaluation Report, the Keynote Address and the 18 unclassified papers, presented at the Information Systems Technology (IST) Panel Symposium held in Aalborg, Denmark from 5th to 7th October 1998.

The papers presented covered the following headings:

- Spectrum Management and Use
- Emerging Technology and Criteria

NORTH ATLANTIC TREATY ORGANIZATION

RESEARCH AND TECHNOLOGY ORGANIZATION
BP 25 • 7 RUE ANCELLE
F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE
Télécopie 0(1)55.61.22.99 • Télex 610 176

DIFFUSION DES PUBLICATIONS
RTO NON CLASSIFIEES

L'Organisation pour la recherche et la technologie de l'OTAN (RTO), détient un stock limité de certaines de ses publications récentes, ainsi que de celles de l'ancien AGARD (Groupe consultatif pour la recherche et les réalisations aérospatiales de l'OTAN). Celles-ci pourront éventuellement être obtenues sous forme de copie papier. Pour de plus amples renseignements concernant l'achat de ces ouvrages, adressez-vous par lettre ou par télécopie à l'adresse indiquée ci-dessus. Veuillez ne pas téléphoner.

Des exemplaires supplémentaires peuvent parfois être obtenus 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 sur la liste d'envoi de l'un de ces centres.

Les publications de la RTO et de l'AGARD sont en vente auprès des agences de vente indiquées ci-dessous, sous forme de photocopie ou de microfiche. Certains originaux peuvent également être obtenus auprès de CASI.

CENTRES DE DIFFUSION NATIONAUX

ALLEMAGNE
Fachinformationszentrum Karlsruhe
D-76344 Eggenstein-Leopoldshafen 2

BELGIQUE
Coordinateur RTO - VSL/RTO
Etat-Major de la Force Aérienne
Quartier Reine Elisabeth
Rue d'Evere, B-1140 Bruxelles

CANADA
Directeur - Gestion de l'information
(Recherche et développement) - DRDGI 3
Ministère de la Défense nationale
Ottawa, Ontario K1A 0K2

DANEMARK
Danish Defence Research Establishment
Ryvangs Allé 1
P.O. Box 2715
DK-2100 Copenhagen Ø

ESPAGNE
INTA (RTO/AGARD Publications)
Carretera de Torrejón a Ajalvir, Pk.4
28850 Torrejón de Ardoz - Madrid

ETATS-UNIS
NASA Center for AeroSpace Information (CASI)
Parkway Center, 7121 Standard Drive
Hanover, MD 21076-1320

FRANCE
O.N.E.R.A. (Direction)
29, Avenue de la Division Leclerc
92322 Châtillon Cedex

GRECE
Hellenic Air Force
Air War College
Scientific and Technical Library
Dekelia Air Force Base
Dekelia, Athens TGA 1010

ISLANDE
Director of Aviation
c/o Flugrad
Reykjavik

ITALIE
Aeronautica Militare
Ufficio Stralcio RTO/AGARD
Aeroporto Pratica di Mare
00040 Pomezia (Roma)

LUXEMBOURG
Voir Belgique

NORVEGE
Norwegian Defence Research Establishment
Attn: Biblioteket
P.O. Box 25
N-2007 Kjeller

PAYS-BAS
RTO Coordination Office
National Aerospace Laboratory NLR
P.O. Box 90502
1006 BM Amsterdam

PORTUGAL
Estado Maior da Força Aérea
SDFA - Centro de Documentação
Alfragide
P-2720 Amadora

ROYAUME-UNI
Defence Research Information Centre
Kentigern House
65 Brown Street
Glasgow G2 8EX

TURQUIE
Millî Savunma Başkanliği (MSB)
ARGE Dairesi Başkanliği (MSB)
06650 Bakanliklar - Ankara

AGENCES DE VENTE

NASA Center for AeroSpace Information (CASI)
Parkway Center
7121 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
Document Delivery,
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)

Imprimé par le Groupe Communication Canada Inc.
(membre de la Corporation St-Joseph)
45, boul. Sacré-Cœur, Hull (Québec), Canada K1A 0S7

NORTH ATLANTIC TREATY ORGANIZATION

RESEARCH AND TECHNOLOGY ORGANIZATION

BP 25 • 7 RUE ANCELLE

F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE

Telefax 0(1)55.61.22.99 • Telex 610 176

DISTRIBUTION OF UNCLASSIFIED RTO PUBLICATIONS

NATO's Research and Technology Organization (RTO) holds limited quantities of some of its recent publications and those of the former AGARD (Advisory Group for Aerospace Research & Development of NATO), and these may be available for purchase in hard copy form. For more information, write or send a telefax to the address given above. **Please do not telephone**.

Further copies are sometimes available from the National Distribution Centres listed below. If you wish to receive all RTO publications, 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 publications may be purchased from the Sales Agencies listed below, in photocopy or microfiche form. Original copies of some publications may be available from CASI.

NATIONAL DISTRIBUTION CENTRES

BELGIUM
Coordinateur RTO - VSL/RTO
Etat-Major de la Force Aérienne
Quartier Reine Elisabeth
Rue d'Evere, B-1140 Bruxelles

CANADA
Director Research & Development
Information Management - DRDIM 3
Dept of National Defence
Ottawa, Ontario K1A 0K2

DENMARK
Danish Defence Research Establishment
Ryvangs Allé 1
P.O. Box 2715
DK-2100 Copenhagen Ø

FRANCE
O.N.E.R.A. (Direction)
29 Avenue de la Division Leclerc
92322 Châtillon Cedex

GERMANY
Fachinformationszentrum Karlsruhe
D-76344 Eggenstein-Leopoldshafen 2

GREECE
Hellenic Air Force
Air War College
Scientific and Technical Library
Dekelia Air Force Base
Dekelia, Athens TGA 1010

ICELAND
Director of Aviation
c/o Flugrad
Reykjavik

ITALY
Aeronautica Militare
Ufficio Stralcio RTO/AGARD
Aeroporto Pratica di Mare
00040 Pomezia (Roma)

LUXEMBOURG
See Belgium

NETHERLANDS
RTO Coordination Office
National Aerospace Laboratory, NLR
P.O. Box 90502
1006 BM Amsterdam

NORWAY
Norwegian Defence Research Establishment
Attn: Biblioteket
P.O. Box 25
N-2007 Kjeller

PORTUGAL
Estado Maior da Força Aérea
SDFA - Centro de Documentação
Alfragide
P-2720 Amadora

SPAIN
INTA (RTO/AGARD Publications)
Carretera de Torrejón a Ajalvir, Pk.4
28850 Torrejón de Ardoz - Madrid

TURKEY
Millî Savunma Başkanliği (MSB)
ARGE Dairesi Başkanliği (MSB)
06650 Bakanliklar - Ankara

UNITED KINGDOM
Defence Research Information Centre
Kentigern House
65 Brown Street
Glasgow G2 8EX

UNITED STATES
NASA Center for AeroSpace Information (CASI)
Parkway Center, 7121 Standard Drive
Hanover, MD 21076-1320

SALES AGENCIES

NASA Center for AeroSpace Information (CASI)
Parkway Center
7121 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
Document Delivery,
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 22161
United States
(also available online in the NTIS Bibliographic Database or on CD-ROM)

Printed by Canada Communication Group Inc.
(A St. Joseph Corporation Company)
45 Sacré-Cœur Blvd., Hull (Québec), Canada K1A 0S7

ISBN 92-837-0003-1