Contents

Cover: The Széchenyi Thermal Baths, Budapest (Foto: Angela Gortva)

Foreword

szabo@hit.bme.hu

Our journal is continuing with the practice of publishing English issues regularly, currently twice a year, in July and in January. The present issue contains English versions of reviewed research papers, carefully selected from the preceeding five Hungarian issues. In general, we also consider papers from open call, therefore the editors would like to encourage prospective authors to submit their results specifically for the English issues. Being a selection, the papers' topics span a wide range of issues of current interest as the reader can see from the short summaries below. We present the papers in the order of their original publication times in the respective Hungarian issues in February through June 2008.

László Toka and Attila Vidács in their paper "Managing users in a peer-to-peer storage system", deal with incentives for peers so that they contribute to the service quality. Two different approaches are suggested: either each peer's use of the service be limited to his/her contribution level (symmetric schemes), or that storage space be bought from and sold to peers by a system operator that seeks to maximize profit. Using a non-cooperative game model to take into account user selfishness, these mechanisms are studies with respect to the social welfare performance measure and necessary and sufficient conditions are given for one scheme to socially outperform the other.

László Muka and Gábor Lencse developed a so-called meta-methodology (MM) for efficient simulation of infocommunication systems and related processes. The goal of this development is to support the use of the most efficient method in any phase of the simulation process. The authors identify factors that influence the simulation problem contexts and make them dynamic, then formulate the requirements on the MM determined by the dynamic simulation problem contexts taking into account the issue of efficiency and also that the simulation method itself is a hard-system method. On this basis the methodology set of MM is defined as a set of hard- and soft-system methods appropriate for different simulation problem contexts.

Lóránt Farkas, Lajos Nagy and Andrea Farkasvölgyi in their paper "Wave Propagation Channel Simulation by Satellite-to-Indoor Radio Link" investigate the propagation characteristics of the satellite-to-indoor propagation channel. First they examine how the polarization state of a complex harmonic field can be described, and then the results of the first simulations of the polarization state are presented. A modified 3D ray-launching tool is utilized for the coverage prediction. The dependence of the indoor wave on the elevation angle of the satellite and the wideband characteristics of the channel are analyzed: the delay spread characteristics and Doppler spread, caused by satellite movement. The applicability of MIMO systems in satellite communications are also dealt with.

An *international group of authors led by Csaba Ferencz* of Lóránd Eötvös University of Sciences, Budapest, Hungary that also includes researchers from Lviv, Ukraine and Troick, Russia reports on an interesting space investigation project. Their SAS2 ULF-VLF electromagnetic wave analyzer system successfully operated on board of the Compass-2 satellite launched in 2006. The measuring systems and instruments worked well and the authors found some interesting phenomena in the observed ULF-VLF electromagnetic database detected by SAS-K2: whistler-doublets (observed earlier in 1989, by the first version of SAS on board of IK-24) and "spiky" whistlers. Further signals, propagated in ducted whistler mode between two inhomogeneous surfaces were successfully identified first time during this mission.

Bálint Tóth and Géza Németh describe the first application of the Hidden-Markov-Model based text-to-speech synthesis method to the Hungarian language. The HMM synthesis has numerous favorable features: it can produce high quality speech from a small database, theoretically it allows to change characteristics and style of the speaker and emotion expression may be trained with the system as well.

György Szaszák and Klára Vicsi in their paper "Using Prosody for the Improvement of Automatic Speech Recognition" describe sentence, phrase and word boundary detection based on prosodic features, implemented in a HMM-based prosodic segmentation tool. Integrated into a speech recognizer, an N-best rescoring is performed based on the output of the prosodic segmenter, which determines the prosodic structure of the utterance. In an ultrasonography task, 3.82% speech recognition error reduction was obtained using a simplified bigram language model.

István Pintér presents a novel speech enhancement method which is based on the concepts of reconstructed phase-space and dimension embedding. The proposed

algorithm separates the speech from noise using a nonlinear transformation in a transformed domain. Recent results in case of uncorrelated, additive noise are presented in this paper.

Szilárd Zsigmond, Marcell Perényi and Tibor Cinkler in their paper "OSNR Based Routing in WDM Optical Networks" propose new physical impairments based routing and wavelength assignment (ICBR) methods where the control plane has influence on the signal power of the Wavelength Division Multiplexed (WDM) channels in metro-optical networks. They give the exact integer linear programming (ILP) formulation of the method. The proposed algorithm can be used in existing WDM optical networks where the nodes support signal power tuning and it outperforms the traditional existing schemes.

The paper by *Áron Szabó and Szilárd Zsigmond* "Determining the Optimal Signal Power Based on Physical Effects in CWDM Optical Networks" presents an analytical model and calculation results for the signal quality degradation in a 8-channel and 18-channel, 2.5 Gbps coarse wavelength-division-multiplexing (CWDM) system. Based on the proposed model and performed analysis the optimal value of the signal power at the transmitter point is given. Modeling of chromatic dispersion and Raman-scattering, the two main constraints of CWDM optical networks is also presented in detail.

This issue will be published a short time before a major technical event, the *Networks 2008* conference will be held in Budapest. This conference is one of the most significant events in telecommunication planning worldwide, has a long tradition since the idea of organizing this kind of event was born within the International Telecommunication Union, then CCITT, in the late seventies, and Budapest is the only city so far that hosts it second time. We hope that our readers will enjoy the history of the Networks conference-series by *Prof. György Lajtha,* one of the founding members of the organizing committee.

Finally, we included short descriptions of two large-scale European Community research projects.

ISIS was a "Network of Excellence" project supported by the FP6 Programme of the European Union. The project acronym stands for "Infrastructures for broadband access in wireless/photonics and Integration of Strength in Europe". ISIS integrated the research activities of 19 organisations from 12 different countries and aimed at strengthening European scientific and technological excellence in low cost optical solutions for broadband access, and the process of merging of wireless and photonic technologies.

MUSE (Multi Service Access Everywhere) was also a European IST FP6 "Integrated Project" during 2004-2007 recognized the leading role of the Ethernet-based technologies in the access networks in serving the growth of the number of subscribers and the magnitude of access services. The project aimed at developing new and unified solutions which are able to serve tens of thousands of users under favorable, low costs. To carry out this task, several large European telecommunication companies including Alcatel-Lucent, Ericsson, Siemens, Deutsche Telekom, British Telecom, France Telecom and Poland Telecom as well as leading research organizations took part in the project.

László Zombory
Chairman of
Editorial Board

Csaba A. Szabó
Editor-in-Chief

Managing users in a peer-to-peer storage system

LÁSZLÓ TOKA, ATTILA VIDÁCS

High Speed Networks Laboratory, Department of Telecommunications and Media Informatics, Budapest University of Technology and Economics

{toka, vidacs}@tmit.bme.hu

Keywords: peer-to-peer networks, storage, backup, game theory, incentives, pricing

When managing a peer-to-peer storage system, where users store their data on the disks of their peers for security, availability and accessibility reasons, the lack of incentives for peers to contribute to the service calls for suitable solutions. We suggest two different approaches: either each peer's use of the service be limited to her contribution level (symmetric schemes), or that storage space be bought from and sold to peers by a system operator that seeks to maximize profit. Using a non-cooperative game model to take into account user selfishness, we study these mechanisms with respect to the social welfare performance measure, and give necessary and sufficient conditions for one scheme to socially outperform the other.

1. Introduction

Along with the convergence in every field of the information technology, various kinds of digital content are now likely to be created and accessed from several types of devices connected to different networks. Therefore, an appropriate system for storing users' data should offer a wide range of services, such as ease of access, protection against device failures, versioning and short transfer time from the system to a given device. In this context, the possibility of storing data online appears as a promising solution.

To exploit the great opportunity, many companies propose online data storage service, most of them offering a given storage capacity (up to 25 GB) for free, with the possibility of extending that quota to a higher value for a fixed price (1$/GB/year). However, while running such a storage service implies owning huge storage capacities and affording the associated energy and warehouse costs, one can imagine using smaller but numerous storage spaces of the users themselves.

In a peer-to-peer (p2p) storage system, the participants are at the same time the providers and the users of the service: each participant offers a part of her disk space to provide the service to the others, and benefits from storing her data onto the system. The added value of the service then comes from the protection against failures provided by the system, from the ease of data access, from the versioning management that may be included, and from the difference in the amount of data stored into the system versus offered for service.

An online storage service is valuable only if data availability is assured, therefore to cope with disk failures and with participants disconnecting their disks from the system, data replicates must be spread over several (sufficiently reliable) peers to guarantee that data are almost always available. A functional p2p storage system needs the participants to offer sufficient fractions of their disk space to the system, and to remain online for long periods. However, both of these requirements imply costs (or at least constraints) for participants, who may be reluctant to devote their resources to the system instead of using them for their own needs. The work presented in this paper focuses on the incentives to make participants contribute to the system. We consider that users behave selfishly, i.e., are only sensitive to the quality of service they experience, regardless of the effects of their actions on the other users, therefore the framework of *Non-Cooperative Game Theory* [7] is particularly well-suited to study the interactions among peers.

While the economic aspects of p2p file sharing networks have already been extensively studied (see [2, 4,8,9] and references therein), there are, to our knowledge, no works tackling the economics of p2p storage networks, although a basic difference exists: in file sharing systems when a peer provides some files to the community, she gives value for *all* users since they all can access the data she proposes; in a p2p storage system, on the other hand, the storage space offered by a user can be shared among different peers but each part is then devoted to only *one* user.

The existing literature on p2p storage systems mainly focuses on security, reliability and technical feasibility issues [3,6,10], whereas the incentive aspect has received little attention. Only solutions that do not imply financial transactions are considered in current works, therefore to create some incentives to participate, the counter payment for providing service is usually the service in question as well. This approach leads to a scheme where every peer should contribute to the system in terms of service at least as much as she benefits from others [5,11]; we call such a mechanism a *symmetric* scheme.

We also investigate solutions based on monetary exchanges: users can "buy" storage space for a fixed unit price, and "sell" their own memory space to the system at another unit price. It is known from economics that

when those unit prices are fixed by the supply and demand curves (as in a perfect market), the selfish user choices lead to a socially efficient situation. However, it is more likely here that the system is managed by a profit-maximizing entity that fixes prices so as to maximize revenue, therefore we study the *profit oriented price-based* scheme.

Our target is to decide which scheme is more convenient for the society. We consider the social welfare, i.e., the total value that the system has for all participants, as the measure of comparison. Under some assumptions on the user utility functions, we derive a necessary and sufficient condition for symmetry-based systems to outperform revenue-oriented management. We obtain that user heterogeneity tends to favor pricing-based schemes that are more flexible, and, above a given user heterogeneity threshold, even a monopoly-managed system will be socially better than a system imposing symmetry.

The paper is organized as follows. The next section presents the user preference model and the two incentive mechanisms mentioned earlier. After that, we define the social welfare performance measure and compute its value for those two schemes. We compare them in the subsequent section to determine the management scheme that is best suited to a given society; and the last section contains our conclusions.

2. Our model

2.1 Data availability, redundancy and transfers

In a p2p storage system the availability of the stored data is considered as the most important factor in user's appreciation. There are no direct means to guarantee that a given user disk storing a specific file will be online 100% of the time, but to ensure data availability, the system can introduce several tools, such as data replication and redundant coding. We suppose here that the system, when detecting that a peer has gone offline, triggers a recovery of the data stored in that peer from the replicas in the system, and a new storage of those data onto other peers. Likewise, when a peer comes back online, then a new data load will be transferred onto her offered storage space, independently what and whose data she was storing before.

Such a data protection mechanism implies data transfers, and therefore non-monetary costs due to resource consumption (CPU, bandwidth utilization, etc.). A peer i is concerned by those data transfers in two situations: when she comes back online after an offline period (receives new data load), and when other peers enter and leave the system (upload traffic if user i stores replicates of the leaving user's data, download traffic when user i has to store more data). The mean data transfer associated with the first situation is thus proportional to the amount of capacity C_i she offers to the system, and to the mean number of online-offline cycles per unit of time: denoting by t_i^{on} (resp. t_i^{off}) the mean duration of online (resp. offline) periods of user i. The corresponding mean amount of transferred data is then proportional to $C_i/(t_i^{\text{on}}+t_i^{\text{off}})$. The mean amount of data transferred to and from user i per unit of time in the second situation is proportional to the weighted (by the offered capacity) mean $\bar{\mu}$ of peer status changes per unit of time. This term appears only at those peers who offer storage space, and only during the time they are online.

Consequently, the transfer cost perceived by user i for offering capacity C_i with the mean availability π_i is expressed by, $C_i\pi_i\left(\delta_i/t_i^{\text{on}}+\gamma_i\bar{\mu}\right)$, where δ_i and γ_i are parameters that reflect the user characteristics such as sensitivity, access bandwidth, or hardware profile.

2.2 User preferences

We describe user preferences by a utility function which reflects the benefit of using the service by storing C_i^s data amount in the system, the cost of offering storage space $C_i^o := \pi_i C_i$ for other users, and the monetary transactions, if any. The utility U_i of user i is of the form

$$U_i\left(C_i^s, C_i, t_i^{\text{on}}, t_i^{\text{off}}, \epsilon_i\right) = V_i(C_i^s) - \underbrace{O_i(C_i\pi_i) - C_i\pi_i\left(\delta_i/t_i^{\text{on}}+\gamma_i\bar{\mu}\right)}_{P_i(C_i, t_i^{\text{on}}, t_i^{\text{off}})} - \epsilon_i,$$

where

- $V_i(C_i^s)$ is user i's valuation of the storage service, i.e., the price she is willing to pay to store the amount C_i^s of data in the system.
- $P_i(C_i, t_i^{\text{on}}, t_i^{\text{off}})$ is the overall non-monetary cost of user i for offering capacity C_i to the system with availability π_i. It consists of two distinct costs:
- the opportunity cost $O_i(C_i\pi_i)$ of offering storage capacity for other users (during online periods) instead of using it for her own needs;
- $C_i\pi_i\left(\delta_i/t_i^{\text{on}}+\gamma_i\bar{\mu}\right)$ data transfer costs due to the data protection mechanism implemented by the system as described in the previous subsection.
- ϵ_i is the monetary price paid by user i for the service taken. This term is 0 in case of a symmetric scheme, and otherwise equals to the price difference between the charge for storing her data into the system and the remuneration for offering her disk space.

2.3 Incentive schemes for cooperation

We assume here that users selfishly choose strategies to maximize their utilities and apart from C_i^s and C_i, each user i can also decide about her behavior related to availability π_i. We describe the two types of incentive mechanisms that we intend to compare later. Both schemes may imply the existence of a central authority or clearance service to supervise peer behavior and/or manage payments: as the model aims to give hints for commercial applications, we do not try to avoid such a centralized system control.

Symmetric schemes

As evoked in the introduction about management solutions without pricing, the principle of those schemes is that users are invited to contribute to, at least as

much as they take from, the other users, i.e., it is imposed to each user *i* that $C_i^o \geq C_i^s$. The availability of the peer is checked (e.g., at randomly chosen times) to ensure that $C_i^o = \pi_i C_i$ exceeds the peer's service use C_i^s.

Payment-based schemes

We consider a simple payment-based mechanism where users can "buy" storage space in the system for a unit price p^s (per byte and per unit of time) and "sell" some of their disk capacity for a (uptime-average) unit price p^o. The (possibly negative) monetary amount that user *i* is charged is then $\epsilon_i = p^s C_i^s - p^o C_i^o$. We assume that the prices are set by a system operator so as to maximize her revenue, knowing *a priori* the reactions of the users. The operator can thus drive the outcome of the game to the most profitable situation for herself, and in this sense, she acts as the leader of a Stackelberg (or leader-follower) game [7]. In a real implementation of the mechanism, the operator may not perfectly know the user reactions, but an iterative groping of prices can converge to the profit-maximizing ones.

2.4 User behavior related to availability

As given before, a user *i* has four strategic variables, namely her offered C_i and stored C_i^s capacities, and her mean online t_i^{on} and offline t_i^{off} period durations. Based on the utility equation, when C_i^s and C_i are fixed, the utility of each user is increasing in t_i^{on}, so t_i^{on} will be set by the selfish user to the reachable maximum value $\bar{t}_i^{on}$, which is constrained by uncontrolled events (power black-out, accidents, hardware failures, etc.) that may force the user off the network. Notice that this selfish decision is profitable to the whole network: longer online periods mean fewer data protection transfers and therefore smaller costs for the system (the parameter $\bar{\mu}$ in the utility equation being small).

Note also that by introducing $p_i^{\min} := \delta_i / \bar{t}_i^{on} + \gamma_i \bar{\mu}$ the transfer costs simply write as $C_i^o p_i^{min}$.

2.5 User supply and demand functions

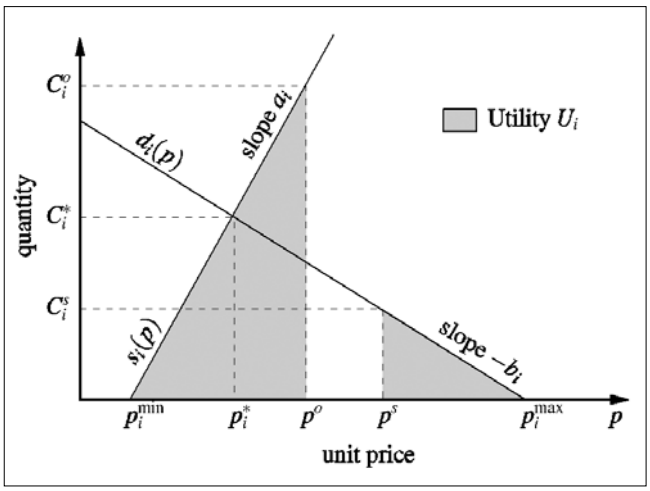
Supply and demand functions are widely used in economics, and are derived from the utility of consumers and cost functions of providers, respectively. For a user *i* we call supply function (resp. demand function) the function $s_i(p)$ (resp. $d_i(p)$) such that for a given $p \geq 0$, $s_i(p)$ (resp. $d_i(p)$) is the amount of storage capacity that user *i* would choose to sell (resp. buy) if she were paid (resp. charged) a unit price *p* for it.

Our illustrative results consider quasi-linear form of affine supply and demand functions and quadratic valuation and opportunity cost functions. Under this consideration, a user *i* is entirely described by four parameters *(see Figure 1)*:

- two price thresholds, namely p_i^{min} and p_i^{max}, that respectively represent the minimum value of the selling unit price and the maximum value of the buying unit price;
- two price sensitivities a_i and b_i, that respectively correspond to the increase of sold capacity when selling price grows and the decrease of bought storage space along with the growth of the buying unit price.

The total supply function $S := \sum_{i \in \mathcal{I}} s_i$ is then a (piecewise affine) increasing convex function on the interval $[\min_i p_i^{min}, \max_i p_i^{min}]$. Likewise, the total demand function $D := \sum_{i \in \mathcal{I}} d_i$ is decreasing and convex on $[\min_i p_i^{max}, \max_i p_i^{max}]$, as illustrated in *Figure 2*.

Figure 1. User reactions to prices and user's valuation



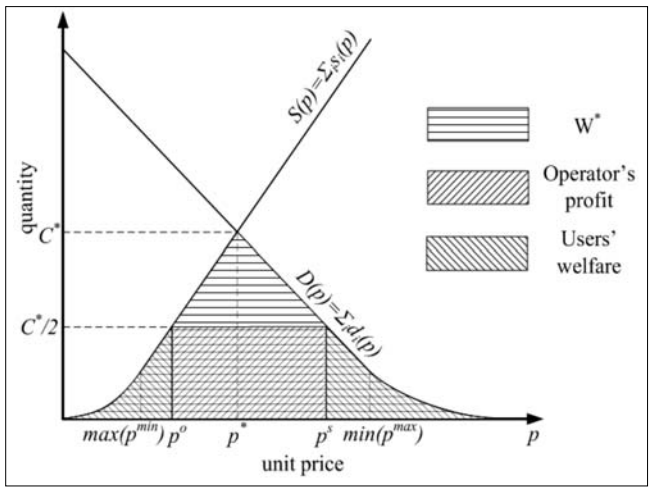
3. Performance of incentive mechanisms

We call the performance measure, used in the following to compare incentive schemes, (social) welfare and denote the sum of the utilities of all agents in the system by *W:*

$$W := \sum_{i \in \mathcal{I}} V_i(C_i^s) - P_i(C_i^o).$$

Notice that no prices appear in the social welfare expression, since all system agents are included, the operator as well that receives or gives payments, if any, and whose utility is her revenue.

Figure 2.
Total supply S and demand D functions, maximum social welfare (lined zone) and surplus repartition (hatched zones) in the case of revenue-driven monopoly under Assumption 1



3.1 Optimal value of social welfare

The optimal situation that the system can attain corresponds to a maximization problem of *W*, subject to the feasibility constraint $\sum_i C_i^o \geq \sum_i C_i^s$; it can be solved by the Lagrangian method. The maximal value *W** as well as the so-called "shadow price" *p** are illustrated in *Figure 2*. Remark that this optimal situation is reached with a special payment-based scheme where $p^o = p^s = p^*$.

3.2 Performance of symmetric schemes

Under a symmetry-based management scheme, each user *i* chooses C_i^o and C_i^s so as to maximize her own valuation, respecting the $C_i^o \geq C_i^s$ constraint.

In [13] it is shown that user *i* maximizes her utility at the point $C_i^s = C_i^o = C_i^*$, as illustrated in Figure 1: this corresponds to the case where every user "exchanges" capacity at the *virtual unit price* p_i^*. However, compared to the socially optimal situation, each and every user loses a part of her utility, thus the system is suboptimal (in terms of social welfare). The loss of welfare depends on the heterogeneity of the users' p_i^*s; if all users have the same p_i^*, then symmetric management schemes maximize social welfare.

3.3 Performance of profit-oriented pricing schemes

When a profit-driven monopoly is employed, the system operator strives to extract the maximum profit out of the business by tuning the prices p^s and p^o.

Figure 2 plots the total supply (*S*) and the total demand (*D*) as functions of the unit selling price p^o and the unit buying price p^s, respectively. Our first remark is that p^o and p^s will be chosen such that the demand would be *exactly* satisfied by the supply: otherwise it is always possible for the operator to decrease p^o (if oversupply) or increase p^s (if overdemand) to strictly improve its revenue. The operator revenue with such prices is then the area of the rectangle displayed in Figure 2, embedded within the triangle-shape zone whose area is the maximum value of social welfare.

To be able to predict the maximal profit generating strategy of the monopoly, we make the following assumption regarding user price thresholds.

Assumption 1

The repartitions of price thresholds $p_{\min}$ and $p_{\max}$ *are such that their variances on the user set are below a certain level* (for details, see [13]).

Moreover, user profile values a_i *(resp.* b_i*) are independent and identically distributed, and each user's* a_i *and* b_i *are independent.*

Proposition 1

Under Assumption 1, a profit-oriented pricing yields the social welfare W_{mon} *such that*

$$W^* - W_{mon} = \frac{1}{8} C^{*2} \left(\frac{1}{\sum_i a_i} + \frac{1}{\sum_i b_i} \right)$$

4. Which management to prefer?

When we compare the outcomes of the two schemes, we immediately have the following result.

Proposition 2

Under Assumption 1, symmetric schemes socially outperform profit-oriented pricing mechanisms if and only if $\frac{1}{4} C^{*2} \left(\frac{1}{\sum_i a_i} + \frac{1}{\sum_i b_i} \right) \geq \sum_i (a_i + b_i) (p^* - p_i^*)^2$.

In other words, if the global shadow price and the users' virtual prices are alike, the symmetric scheme reaches higher social welfare.

Proposition 2 combines the four user heterogeneity factors, namely the price thresholds $p_{\min}$, $p_{\max}$ and the price sensitivities *a, b*, to determine the best mechanism in terms of social welfare. Whereas the right-hand term is the (weighted) variance of the p_i^*, the left-hand term is hard to interpret.

We thus suggest to take a look at the particular cases where user heterogeneity lies entirely on price sensitivities (resp. on price thresholds).

4.1 Homogeneous price thresholds

We consider here that users only differ by their price sensitivities a_i and b_i, and they have the same price thresholds $p_i^{\min}$ and $p_i^{\max}$. This case has been studied in a previous work, we therefore recall the main results and refer the interested reader to [12] for details: it is proven that

$$W_{sym} = \left(\frac{1}{\sum_i a_i} + \frac{1}{\sum_i b_i} \right) \sum_i \left[\frac{1}{\frac{1}{a_i} + \frac{1}{b_i}} \right] W^*,$$

$$W_{mon} = \frac{3}{4} W^*,$$

which yields the following comparison.

Proposition 3

Under our assumptions, symmetric schemes socially outperform profit-oriented pricing mechanisms if and only if $\left(\frac{1}{\sum_i a_i} + \frac{1}{\sum_i b_i} \right) \sum_i \frac{1}{\frac{1}{a_i} + \frac{1}{b_i}} \geq \frac{3}{4}$.

Moreover, if the couples (a_i, b_i) *are independently chosen for all users and identically distributed, then the inequality holds if and only if*

$$\frac{\mathbb{E}[f(a,b)]}{f(\mathbb{E}[a], \mathbb{E}[b])} \geq \frac{3}{4}$$

when the number of users tends to infinity (law of large numbers), with $f : (x, y) \mapsto \frac{1}{1/x + 1/y}$.

Since the function *f* is strictly concave, from Jensen's inequality the left-hand term is always smaller than 1, and decreases as the dispersion of (*a, b*) increases. Remark that when (*a, b*) are deterministic then the left-hand term equals to 1 and symmetric schemes are better than profit-oriented ones, as we remarked above.

Let us have a look at Proposition 3 for two simple examples of (*a, b*)'s distribution, assuming that *a* and *b* are independent variables.

Uniform distribution:

If *a* (resp. *b*) is uniformly distributed over $[0, a_{\max}]$ (resp. $[0, b_{\max}]$), the inequality always holds.

Exponential distribution:

If *a* (resp. *b*) follows exponential distribution with parameter μ_a (resp. μ_b), either a symmetric or a profit-oriented mechanism is socially preferable depending on the relative values of μ_a and μ_b.

4.2 Homogeneous price sensitivities

We now consider the case where the price thresholds p_i^{min} and p_i^{max} are user-specific, but the price sensitivities a_i and b_i are identical for every user. The couples (p_i^{min}, p_i^{max}) are independent and identically distributed among users; moreover p_i^{min} is independent of p_i^{max} for all users.

Proposition 4

Under these assumptions, managing mechanisms based on symmetry are always socially better (in terms of social welfare) than profit-oriented pricing mechanisms.

5. Conclusion

In this work we have addressed the problem of user incentives in a peer-to-peer storage system. Using a game theoretical model to describe selfish reactions of all systems actors (users and the operator), we have studied and compared the outcomes of two possible managing schemes, namely symmetry-based and profit-oriented payment-based policies.

Not only the size of the offered storage space was targeted with incentives, but as the availability and reliability are particularly important issues in storage systems, the churn as well. By comparing the social welfare in the two cases, under some assumptions on user preferences, we exhibited a necessary and sufficient condition for a type of management to be preferable to the other: it appears that profit oriented payment-based schemes may be socially better than symmetric ones under some specific circumstances, namely if the heterogeneity among user profiles is high.

Acknowledgement

László Toka hereby thanks Patrick Maillé for his continuous and extremely helpful support during his work, and to the HTE Education Committee for considering his MSc thesis valuable to reward and publish.

Authors

ATTILA VIDÁCS was born in Budapest, in 1973. He received the MSc and PhD Degrees from the Budapest University of Technology and Economics at the Department of Telecommunications and Media Informatics, Faculty of Electrical Engineering and Informatics, in 1996 and 2001, respectively. During 1997, he worked as a visiting researcher at the Research and Development Center of the Nippon Telegraph and Telephone Corp., Tokyo, Japan. Currently he is Associate Professor at the Dept. of Telecommunications and Media Informatics, Budapest University of Technology and Economics. His research interests are in the field of teletraffic modeling and traffic engineering, sensor networking, and dynamic spectrum access networks.

LÁSZLÓ TOKA graduated in 2007 at Budapest University of Technology and Economics and received his MSc degree in Telecommunications on the Faculty of Electrical Engineering and Informatics. He spent the last two years of his study in France and obtained the engineer diploma of Telecom Bretagne and Eurecom. He also participated in the pre-doctoral education courses of the Networks and Distributed Systems Department at the University of Nice Sophia-Antipolis and received a research master degree. During his undergraduate years László participated in research projects focusing on economic modeling of distributed IT systems and networks. After graduating he followed up on his research domain and enrolled as a PhD candidate at Telecom Paris and at the Budapest University of Technology and Economics.

References

[1] E. Adar and B. Huberman, Free riding on gnutella. Tech. Rep., Xerox parc, 2000.

[2] P. Antoniadis, C. Courcoubetis, R. Mason, Comparing economic incentives in peer-to-peer networks. Computer Networks, 46(1):133–146, 2004.

[3] C. Batten, K. Barr, A. Saraf, S. Treptin, pStore: A secure peer-to-peer backup system. Tech. Rep., MIT Lab. for Computer Science, 2001.

[4] C. Courcoubetis and R. Weber, Incentives for large peer-to-peer systems. IEEE JSAC, 24(5):1034–1050, 2006.

[5] L. Cox and B. Noble, Samsara: Honor among thieves in peer-to-peer storage. In Proc. of 19th ACM Symposium on Operating Systems Principles (SOSP'03), 2003.

[6] P. Druschel and A. Rowstron, PAST: A large-scale, persistent peer-to-peer storage utility. In HotOS VIII, pp.75–80., 2001.

[7] D. Fudenberg and J. Tirole, Game Theory. MIT Press, 1991.

[8] P. Golle, K. Leyton-Brown, I. Mironov, M. Lillibridge, Incentives for sharing in peer-to-peer networks. In Proc. of 3rd ACM Conf. on Electronic Commerce (EC'01), pp.264–267., 2001.

[9] K. Lai, M. Feldman, I. Stoica, J. Chuang, Incentives for cooperation in peer-to-peer networks. In Proc. of 1st Workshop on Economics of Peer-to-Peer Systems, 2003.

[10] M. Lillibridge, S. Elnikety, A. Birrell, M. Burrows, M. Isard A cooperative internet backup scheme. In Proc. of 1st Workshop on Economics of Peer-to-Peer Systems, 2003.

[11] B. Stefansson, A. Thodis, A. Ghodsi, S. Haridi, MyriadStore. Tech. Rep., Swedish Inst. of Comp. Science, 2006.

[12] L. Toka, P. Maillé, Managing a Peer-to-Peer Backup System: Does Imposed Fairness Socially Outperform a Revenue-Driven Monopoly? In LNCS Proc. of 4th International Workshop on Grid Economics and Business Models (GECON), 2007.

[13] L. Toka, Peer-to-peer sharing of hard disks: study of incentive mechanisms, MSc thesis, 2007.

Developing a meta-methodology for efficient simulation of infocommunication systems and related processes

LÁSZLÓ MUKA
Elassys Consulting Ltd., muka.laszlo@elassys.hu
GÁBOR LENCSE
Department of Telecommunications, Széchenyi István University, lencse@sze.hu

Keywords: simulation meta-methodology, dynamic simulation problem context, ICT, hard and soft-systems method, BP

The efficiency of simulation projects aimed at supporting the design of Information and Communication Technology (ICT) and related Business Process (BP) systems in an organisation is influenced by some key factors. The goal of the development of our simulation meta-methodology (MM) is to support the use of the most efficient method in any phase of the simulation process. In this paper we indentify the factors influencing the simulation problem contexts and making them dynamic, then we formulate the requirements on the MM determined by the dynamic simulation problem contexts taking into account the issue of efficiency and also that the simulation method itself is a hard-system method. On this basis we define the methodology set of MM that is a set of hard- and soft-system methods appropriate for different simulation problem contexts. We examine the important features of the MM methodology components, we describe the general features of the simulation methodology (SM) in detail, we propose and also define further requirements on SM determining extra features. We introduce the cycles and the process of MM including alternating way of work and the methodology chains which make MM suitable for dynamic simulation problem contexts.

1. Introduction

The efficiency of simulation projects aimed at supporting the design of ICT and related BP systems in an organisation is influenced by some key factors including also methodological factors. In our earlier papers we have already examined many of these factors and we also investigated the ways of increasing the efficiency [16-21].

It is important to note that in order to improve efficiency of simulation the MM under development focuses not only on the question of direct *efficiency* but also addresses the problems of the *efficacy* and *effectiveness* [9], either by means of first of all soft-system methods and preliminary modelling.

In this paper, first we outline the system focus of application of the meta-methodology and define the process of simulation. We use a new approach: the concept of *the dynamic simulation problem contexts.* We identify the factors influencing simulation problem contexts that is factors influencing simple-complex and unitary-pluralist features and making them dynamic are identified, which are also responsible for the existence of complex-pluralist problem contexts. On this basis we formulate the requirements on the new meta-methodology.

Then, we examine the set of elements of the simulation meta-methodology. As the starting point of formulation of SM, we examine the evolution of the traditional simulation methodologies. We introduce the general features of the proposed simulation methodology and also the new requirements on the SM which we define as special features of SM. We present a brief evaluation of the selection of both SSM (Soft Systems Methodology) and MCM (Modified Conceptual Modelling) methods. In the section about the further elements, we mention TFA (Traffic Flow Analysis) and EFA (Entity Flow-phase Analysis) methods which are proposed for rapid preliminary modelling, and we briefly describe meta-methodology element "goal reduction and linking". We introduce important new elements: the *alternating way of work* of simulation meta-methodology and the *methodology chains* formed by the problem context sequences.

Then, the requirements, which are determined by the dynamic simulation problem contexts, on simulation meta-methodology (MM) are formulated from the point of view of efficiency, taking also into account, that simulation method itself is a hard-systems approach.

On this basis, a set of hard and soft systems methods for MM is defined, which is appropriate for different simulation problem contexts.

Important features of methodology elements of MM are introduced. These elements, which have already been described in our previous papers, are as follows: the typical synthesised Simulation Methodology (SM) with added special features, the Modified Conceptual Models (MCM) methodology, and other methods. The Soft Systems Methodology (SSM) is also presented as the basic soft-systems approach for MM.

The phases, the cycles, and the process of MM (including the alternating way of work and the methodology chains) – which make MM suitable for dynamic simulation problem contexts – are described.

Finally, the functioning of MM in a collaborative modelling environment is examined, which is a frequent situation.

2. Simulation and the environment of simulation

2.1 System Scope of the Simulation Meta-methodology

In this paper we develop a simulation meta-methodology appropriate for the examination of info-communication systems and connected processes.

The system scope of the simulation meta-methodology may be defined by the group of ICT (Information and Communications Technology) and related BP (Business Process) or OP (Organisational Process) systems. ICT and *connected* BP or OP systems form EIS (Enterprise Information Systems) or respectively OIS (Organisational Information Systems).

2.2 Process of Simulation

Definitions of simulation have been proposed by many authors (see for example [25]). Now, for the meta-methodology development purposes we propose the following approaches to the simulation:

Simulation is a process of developing simulation model of the system of interest and performing experiments with the model in order to reach the defined goals.

The *process of simulation* lasts from the identification and investigation of the need for developing a simulation model of a system of interest to providing support to implement results of simulation [15].

In an *organisational environment,* we may look at the process of simulation performed as a *project process,* initiated to reach pre-defined goals, within time and cost limits and with the required quality, and using the assigned resources.

2.3 Dynamic Simulation Problem Contexts

Modelling projects often start with an *unstructured problem situation:* even if there was a consensus about the application of simulation it may turn out in the "Defining Goals" phase that there is no agreement about the questions to be answered [22].

It is often necessary to use the simulation methodology in a soft-systems environment: even the problem structuring ("Defining Goals" phase) may lead to *complex-pluralist problem contexts for simulation* which require the application of a soft-systems approach but the *simulation is a hard-systems approach* appropriate for simple-unitary problem contexts (the problem contexts are described in [11], the features of hard-systems and soft-systems approaches can be found in [8]. Moreover, it is important to remark that the simulation problem context may change *dynamically* in any phase of the simulation process.

Now, we examine the factors influencing the *simulation problem context* according to the *simple-complex* and *unitary-pluralist* dimensions, which make problem contexts often complex-pluralist.

Factors influencing the simple-complex dimension:

- Systems are often only partially observable (for example data are not collected or cannot be collected because of technical reasons or because data sources are located in other systems).
- The systems of interest cannot be easily defined (for example, systems' boundaries are not observable because of data availability problems).
- Simulated systems are of probabilistic nature and may have active parts with independent objectives (for example people in the system may act in opposition to simulation project goals).
- The complexity may increase by taking into account the influences on other systems.

Factors influencing the unitary-pluralist dimension:

- Simulation project is performed in an environment formed by many participants:
  Decision makers, problem solvers (users, analysts, modellers, etc., who may also be decision makers in different phases), whose' views on the world influence the simulation problem context.
- The initial problem structuring often leads to a pluralist set of opinions regarding the goals [22].
- A disagreement can also occur regarding the implementation of results (for example, who is responsible for what during the implementation [22].

Simulation is an efficient method if it used as a hard-systems approach to the problems of simple-unitary contexts, therefore, to be efficient, we should have a *set of methods* appropriate for different contexts and we also should have a *formalised process, a simulation meta-methodology* to control the use of methodologies in dynamic simulation problem contexts.

3. Defining components of the simulation meta-methodology

The set of methods of the simulation methodology should contain a *traditional* simulation methodology (hard-systems method), a method appropriate for problem contexts requiring *soft-systems* approach and also a method *connecting* the hard-systems and soft-systems levels. It is also useful to have methods making the coverage of the simulation process complete supporting the improvement of the efficiency of simulation. In the following, we examine and introduce these elements of the set of methods.

3.1 Synthesis of a Traditional Simulation Methodology with Extra Features

Evaluation of Traditional Simulation Methodologies

The simulation method containing a series of phases has already been described by many authors [1-3,7,26]. These phases represent the highest level of development and application of the simulation model. This description level of the simulation process remains constant regardless of the type of the problem and the objective of the simulation analysis [7]. Furthermore, simulation models can capture the behaviour of both human and technical resources in the system [26]. Examining the methodologies described by the aforementioned

authors, an *evolution* of methodologies may be observed, starting from the *problem-solution-type*, strictly hard approach to the present days' more *soft-approaches*.

The current state-of-the-art can be summarised according to the 3 main stages of methodologies:

Prior-to-modelling stage:

Simulation is project-based: it is a process with pre-defined objectives, which should be reached within a time and cost limit and with the required quality, using the resources assigned to the process. This view shows the collaborative character of a simulation project.

Modelling and experimentation stage:

For different tasks, there is a wide variety of simulation tools, with different model building and experimenting features, therefore methodologies can contain tool-specific features.

After-modelling stage:

Simulation becomes a decision support tool: the outputs of simulation can be regarded as *understanding-type results* supporting decision making rather than *solution-type results* providing an exact solution to a problem. The results of simulation are also *project-type results:* a report should be generated and documented for the defined participants of the project.

Typical Simulation Methodology

As an element of the simulation meta-methodology (MM) we describe a typical hard simulation methodology (SM) comprising six steps (the detailed description of SM is in [20]).

It is not a novel methodology, it is rather a *synthesis* based on the conclusions of the analysis described in the previous section, but we pay special attention to some requirements and define *extra features* for our typical SM.

The six-step process of simulation methodology (in *Figure 1*) has the following phases:

SM1: Defining Goals
SM2: Gathering and Analysing Data
SM3: Model Design and Model Building
SM4: Performing simulation
SM5: Analyzing Results
SM6: Supporting Implementation

Summary of Features of SM

Extra features:

- An *output is defined to each phase* in order to support methodological communication.
- Special attention is paid to *preliminary modelling*.
- Simulation is assigned to support implementation. Decisions in order to avoid disagreement about implementation (there are often different views on implementation of results).

General features:

- SM is a tool-independent methodology.
- SM puts equal emphasis on each of the three main phases.

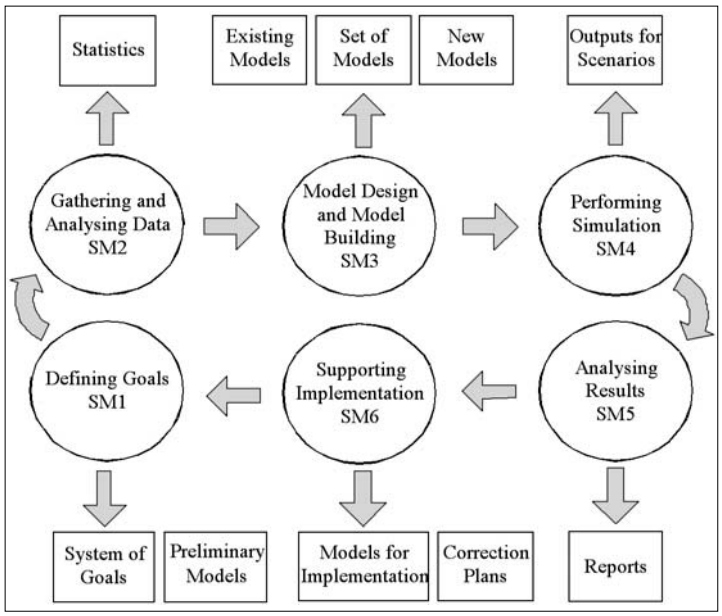


Figure 1. The six-step process of simulation methodology with extra features

- SM can be applied to simulate both BP and ICT elements of an organisational information system.
- SM, like all the examined methodologies, has an *iterative* character, phases or group of the phases can be repeated until they produce a suitable outcome.
- SM has a *cyclic* character, that is, the methodological loop may be closed forming short-cycles or long-cycles:

– There can be any full or partial methodological cycles during a simulation project (short-cycles)
– The simulation models may be reused at any point of time, later, during the life-cycle of the modelled system (long-cycles)

3.2 SSM in the Simulation Meta-Methodology: Short Evaluation of SSM and other Possibilities

SSM is the classic soft-systems approach [8]. Arguments for selecting SSM as MM element may be summarised as follows:

The methodology should be able to face with soft-problem situations both in ICT and BP fields.

The well known approach of UML has the capabilities to face with ICT and BP sides but UML is weak in dealing soft aspects [6]. TSI (Total System Intervention [12]) is rather a framework of methodologies (with a large set of associated methodologies) and there is no known experience of using it in ICT or BP field. For SSM there is a significant amount of applications and experience to use it with or in other methods [10,5].

3.3 MCM in the Simulation Meta-Methodology: Short Evaluation of MCM and other Possibilities

By simulation the dynamic features of systems are investigated, therefore it is necessary to use time in simulation models.

The introduction of time into UML is described in [24], but UML is weak in dealing soft situations as we have already seen. Gregory's method [13, 17] is a method based on SSM and operates with "enhanced" conceptual models but it has no appropriate time tools (synchronisation of model times, time decomposition) which are necessary in a simulation environment and does not differentiate between IT and P systems which is also necessary for efficient simulation.

Usual approaches to use SSM models together with other methods include grafting and embedding [23]. (Examples for grafting and embedding can be found in [4] and [5], respectively).

MCM (SSM with modified conceptual models) may be characterised as an *extension* of SSM models with *extra features* and grafting the methods of using extended models into SSM. This way MCM is applicable both at soft-system and hard-system level, supporting the elimination of the *methodological gap*.

3.4 Further Components

Further elements are the TFA (Traffic Flow Analysis) [17,18] and EFA (Entity Flow-phase Analysis) [16-18] which are methods for rapid preliminary modelling and for goal reduction and linking.

An enterprise has a set of goals with formal and informal features. The goals in a current set of goals influence each other and may also be in conflict with each [14]. Goals of the simulation project should be obtained from higher level goals. The *"SSM problem learning"* method and the *"goal-reduction-linking"* method support the goal setting process of the simulation project.

4. Cycles and working process

4.1 Cycles of the Simulation Meta-methodology

The detailed description of elements, outputs and phases of MM is given in [19] and [20].

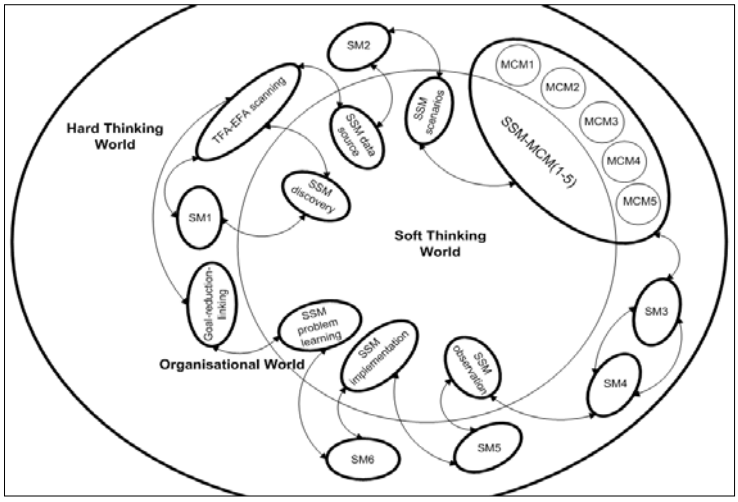


Figure 3.
The alternating way of work of the simulation meta-methodology

The main methodological cycle of MM is the MM1-MM4 cycle (indicated by empty arrows in *Fig. 2*). The progress in the main cycle occurs according to SM1-SM6 steps. In an MM phase there can be usual sub-cycles indicated by dashed lines and arrows. Preliminary modelling may be connected to MM1 or MM2 too and may induce sub-cycles between MM1 and MM2 phases. MCM cycle is shown by dashed lines and an arrow in MM3. It may form its own sub-cycle inside the phase. (A possible sequence of cycles is demonstrated in *Fig. 3*.)

4.2 Working Process of the Simulation Meta-Methodology

In order to be efficient and to be able to address the dynamic problem contexts of simulation we should have a *full and compatible set of methods* covering the whole process of simulation. (This set of methods is introduced in the sections about SM, SSM, MCM and in the "Further Components" section.)

The meta-methodology *governs* the use of the methods during the process of simulation: the meta-methodology supports the use of the suitable method for every situation (simulation problem context) or from other point of view it *directs* the work in the dynamically changing contexts taking into account that the simulation itself is a hard-system method.

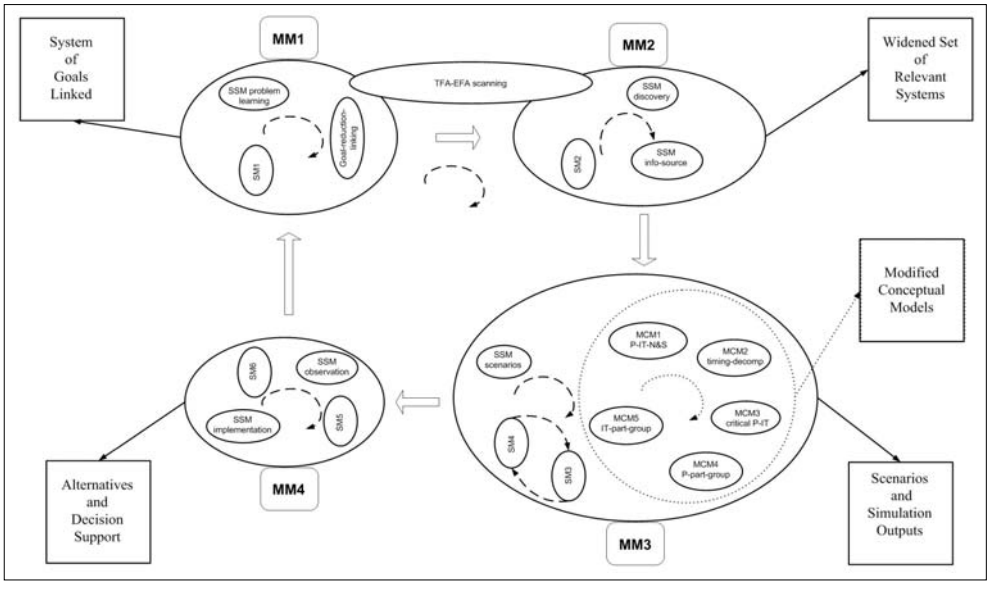


Figure 2.
Elements and cycles of the simulation meta-methodology

In the process of performing simulation (simulation project) usually *dynamic simulation problem contexts* occur.

Therefore MM should have the possibility to "soften up" the methodology and then, after exploring the problem context, to "harden up" again. The *alternating hardening and softening up* the methodology means that after hard cycles (which are directed to find a solution in a given step) it is necessary (or advisable) to use soft cycles, in order to explore the whole situation.

The sequence of hard and soft methods in the process of using the meta-methodology forms a *methodology chain:* in the chain each of the elements (methods) uses the results of the previous element and prepares the use of the next element. The methodology chain is started and finished by a soft method application. The methodology chain may be described by the sequence of simulation problem contexts and by the methods used to the contexts.

Figure 3 shows that Organisational World is divided into two segments: the Hard Thinking World and the Soft Thinking World. Soft-systems methods are situated in the Soft Thinking World and hard-systems methods are in the Hard Thinking World. MCM operates between these two segments. MCM process starts and finishes its operation with the "SSM problem learning" method.

Different methods are connected by a bi-directional connection which indicates that in the process of MM if it is necessary we may re-enter an earlier step. A sequence of steps performed according to connections shows the alternating work of MM. (Of course in the process of operation of MM it may be necessary to use other connections (which are not shown in the figure) between methods.)

5. Summary

In this paper, we presented the further development of the new simulation meta-methodology. Our main goal was to increase the efficiency of simulation by supporting the use of the most efficient method for a given problem context (simulation problem context) in any phase of the simulation process by means of the meta-methodology.

For our examination, we have defined the system scope of the simulation meta-methodology (systems for which we intend to apply the simulation meta-methodology) and we have also defined the process of simulation we used in our considerations.

The factors influencing simulation problem contexts and making them dynamic have been identified.

The requirements on MM determined by the dynamic simulation problem contexts have been described taking into account the point of view of efficiency and also the hard-systems character of the simulation method itself. A set of hard and soft systems methods (appropriate for different simulation problem contexts) for MM has been defined and the most important features of methodology elements of MM have been introduced.

We have given a short overview of the elements of the methodology set of MM, described the general and special features of the typical, synthesised SM and the cycles and the working process of MM (including the alternating way of work appropriate for dynamic simulation problem contexts and the methodology chains).

The important aspects of this paper may be summarised as follows: a complex approach to the efficiency issue of simulation is described (taking into account the whole process of simulation including modelling); on this basis the first formulation of general requirements to the problem is introduced; by developing the simulation meta-methodology (and its methodology elements), an efficient answer to the problem is proposed.

Acknowledgement

We would like to thank Philippe Geril (EUROSIS ETI) for his permission to use our previously published paper "Hard and Soft Approaches in a Simulation Meta-Methodology" [20].

Authors

GÁBOR LENCSE received his M.Sc. in electrical engineering and computer systems at the Technical University of Budapest in 1994 and his Ph.D. in 2000. The area of his research is parallel discrete-event simulation methodology. He is interested in the acceleration of the simulation of info-communication systems. Since 1997, he has been with the Széchenyi István University in Győr, now as Associate Professor. He teaches computer networks and networking protocols. He is founding member of the Multidisciplinary Doctoral School of Engineering, Modelling and Development of Infrastructural Systems at the Széchenyi István University. He has been doing R&D in the field of the simulation of communication systems for the Elassys Consulting Ltd. since 1998. Dr. Lencse has been working part time at the Budapest University of Technology and Economics (the former Technical University of Budapest) since 2005 where he teaches computer architectures.

LÁSZLÓ MUKA graduated in electrical engineering at the Technical University of Lvov in 1976. He got an engineering specialization degree in digital electronics at the Technical University of Budapest in 1981, and obtained the Dr. Univ. degree in architectures of CAD systems in 1987. Dr. Muka finished an MBA at Brunel University of London in 1996. Since 1996 he has been working in the area of simulation modelling of telecommunication systems, including human subsystems. He is a regular invited lecturer in the area of applications of computer simulation for performance analysis of telecommunication systems at the Multidisciplinary Doctoral School of Engineering, Modelling and Development of Infrastructural Systems at the Széchenyi István University of Győr.

References

[1] Churchman, C. W., Ackoff R. L., Arnoff, E. L., "Introduction to Operations Research", John Wiley & Sons, 1957.

[2] Seprődi, L., "A GPSS szimulációs nyelv", Műszaki Könyvkiadó, 1980.

[3] Powis, D., "Understanding Simulation Modeling for the Contact Center", Vanguard Communications Corporation, 2002. http://www.vanguard.net/DicLib_Docs/Simulation_Modeling_dp_0204.pdf

[4] Wilson, B., "Systems: Concepts, Methodologies and Applications", Wiley, Chichester, 1984.

[5] Rodriguez-Ulloa, R., Paucar-Cacers,
"A., Soft System Dynamics Methodology (SSDM):
A Combination of Soft Systems Methodology (SSM)
and System Dynamics (SD)",
In Proc. of 43rd Meeting of the International Society
for the System Sciences, Pacific Grove, CA, 1999.
[6] Al-Humaidan, F., Rossiter, N.,
"Evaluation of System Analysis Methodologies in
a Workflow Context",
InterSymp 2002 – 14th Int. Conference on Systems
Research, Advances in Computer Cybernetics XI,
Ed. by Lasker, G.E., pp.8–13, 2002.
[7] Balachandran, A., Rabuya, L. C.,
Shinde, S., Takalkar, A.
"Introduction to Modeling and Simulation Systems:
Basic Steps and Decisions for Simulation", 2002.
http://www.uh.edu/~lcr3600/simulation/steps.html
[8] Checkland, P.,
"From Optimizing to Learning:
A Development of Systems Thinking for the 1990s",
J. of the Operational Research Society,
Vol.36, No.9, pp.757–767, 1985.
[9] Checkland, P.,
"Soft Systems Methodology in Rational Analysis for
a Problematic World", Ed. by J. Rosenhead,
John Wiley & Sons, 1989.
[10] Curtis, G.,
"Business Information Systems",
Addison-Wesley, Wokingham, UK, 1989.
[11] Jackson, M.C., Keys, P.,
"Towards a System of Systems Methodologies"
J. of the Operational Research Society,
Vol.35, No.6, 1984.
[12] Flood, R. L., Jackson, M, C.,
"Creative Problem Solving - Total Systems Intervention"
John Wiley & Sons, New York, 1991.
[13] Gregory, F.,
"Cause, Effect, Efficiency and Soft Systems Models",
J. of the Operational Research Society,
Vol.44, No.4, 1993.
[14] Koubarakis, M., Plexousakis, D.,
"Business process modelling and design -
a formal model and methodology"
BT Technol. J., Vol.17, No.4, 1999.
[15] Paul, R. J., Hlupic. V., Giaglis, G.,
"Simulation Modelling of Business Processes",
UKAI'98 – UK Academy of Information Systems Conf.,
Lincoln, UK, 1998.
[16] Lencse, G., Muka, L.,
"Expanded Scope of Traffic-Flow Analysis:
Entity Flow-Phase Analysis for Rapid Performance
Evaluation of Enterprise Process Systems"
Proc. of the 2006 European Simulation and
Modelling Conference (ESM'2006), Toulouse, France,
Oct. 23-25, 2006. EUROSIS-ETI, pp.94–98.
[17] Lencse, G., Muka, L.,
"Combination and Interworking of
Four Modelling Methods for Infocommunications
and Business Process Modelling"
Proc. of the 5th Industrial Simulation Conference
(ISC'2007), Delft, The Netherlands, June 11-13, 2007.
EUROSIS-ETI, pp.350–354.
[18] Lencse, G., Muka, L.,
"Investigation of the Spatial Distribution Algorithm of
the Traffic Flow Analysis and of the Entity Phlow-
Phase Analysis"
Proc. of the 2007 European Simulation and Modelling
Conference (ESM'2007), St. Julians, Malta,
Oct. 22-24, 2007. EUROSIS-ETI, pp.574–581.
[19] Muka, L., Lencse, G.,
"Developing a Meta-Methodology Supporting
the Application of Parallel Simulation"
Proc. of the 2006 European Simulation and Modelling
Conference (ESM'2006), Toulouse, France,
Oct. 23-25, 2006. EUROSIS-ETI, pp.117–121.
[20] Muka, L., Lencse,G.,
"Hard and Soft Approaches in
a Simulation Meta-Methodology"
Proc. of the 5th Industrial Simulation Conference
(ISC'2007), Delft, The Netherlands, June 11-13, 2007.
EUROSIS-ETI, pp.17–22, 2007.
[21] Muka, L., Lencse,G.,
"Decision Support Method for Efficient Sequential
and Parallel Simulation:
Time Decomposition in Modified Conceptual Models"
Proc. of the 2007 European Simulation and Modelling
Conference (ESM'2007), St. Julians, Malta,
Oct. 22-24, 2007. EUROSIS-ETI, pp.291–295.
[22] Pidd, M.,
"Operation Research/Management Science Method
in Operations Research in Management",
Ed. by Littlechild, S. and Shutler. M.,
Prentice Hall, UK., 1991.
[23] Rose, J.,
"Information Systems Development
as Action Research - Soft Systems Methodology and
Structuration Theory", Ph.D. Thesis Jeremy Rose,
November 2000.
[24] Hennig, A., Wasgint, R.,
"Performance Modeling of Software Systems in
UML-Tools for the Software Developer",
Proc. of European Simulation Multiconference
(ESM'2002), Darmstadt, Germany, 2002.
[25] Jain, R.,
"The Art of Computer Systems Performance Analysis:
Techniques for Experimental Design, Measurement,
Simulation, and Modeling,"
Wiley–Interscience, New York, April 1991.
[26] Hlupic, V., Robinson, S.
"Business Process Modelling and Analysis Using
Discrete Event Simulation"
Eds. by D. J. Medeiros, E. F. Watson, J. S. Carson
and M. S. Manivannan,
Proc. of the Winter Simulation Conference,1998.

Wave propagation channel simulation by satellite-to-indoor radio link

LÓRÁNT FARKAS, LAJOS NAGY, ANDREA FARKASVÖLGYI

Department of Infocommunication and Electromagnetic Theory, Budapest University of Technology and Economics,

lajos.nagy@mht.bme.hu

Keywords: propagation model, polarization, ray-launching, satellite communication

In our paper we present the simulation of the propagation characteristics of the satellite-to-indoor propagation channel. Our first aim has been to find a correct description of the polarization state of the received inside wave. The result of our first investigations is that the polarization state of the indoor wave signigicantly changes as we move further away from the windows that are the secondary source of radiation. First we examine how the polarization state of a complex harmonic field can be described, and then the results of our first simulations of the polarization state will be presented. A modified 3D ray-launching tool has been utilized for the coverage prediction. A detailed analysis of the dependence of the indoor wave on the elevation angle of the satellite is given, and the wideband characteristics of the channel: delay spread characteristics and Doppler spread, caused by satellite movement are dealt with. The applicability of MIMO systems in satellite communication is also investigated.

1. Introduction

For mobile satellite systems a wide range applications can be foreseen. Indoor operation is one of the key problems in personal wireless communications. Without the possibility of indoor operation, application of satellite systems in personal communications networks would be unlikely. Users of personal communication networks will – sooner or later – force the operators to provide also this kind of service. Therefore, in our opinion, an elaborated tool for the prediction of the indoor penetration would be highly useful.

Due to extreme distances and extreme attenuation the propagation media may be considered a very hostile one. The waves arriving at the buildings do have significantly variable characteristics, but generally for tall buildings it can be assumed that plane waves have some kind of polarization state, mostly an elliptical one.

Once penetrated, the building adds its impact on the wave by multiple reflections, transmission through walls and diffraction through corners and inhomogeneities of building materials.

The field inside the building is a complex one, but it can be assumed to be harmonic and consequently it can be analyzed. As a general framework, we propose a complete simulation tool for the narrow-band and wideband characteristics of the satellite-to-indoor propagation channel.

In this paper we focus as a first step on the polarimetric description of the indoor waves and some conclusions regarding their general polarimetric characteristics in the case of elliptically and linearly polarized incident plane waves will be reached. This first step intends to clarify what kind of antennas would be needed for indoor receivers and what can be achieved using these antennas. The simulation at this stage takes into account multiple reflections and transmissions through walls and is based on a 3D ray-launching tool. Diffraction is intended to be taken into account as a next step.

2. Polarization

Every harmonic vector field can be characterized by its polarization property. Generally speaking we can define polarization as a local property of a harmonic vector field as the curve described by the field strength in a given location. According to [1] the polarization of a radiated wave is “that property of a radiated electromagnetic wave describing the time-varying direction and relative magnitude of the electric field vector; specifically, the figure traced as a function of time by the extremity of the vector at a fixed location in space, and the sense in which it is traced, as observed along the direction of propagation”.

According to [2], polarization may be classified into three categories: linear, circular and elliptical, of which circular is a special case of the elliptical (in fact one extreme of it) and linear is also a special case of elliptical (the other extreme). Polarization in general, except the linear polarization, can be clockwise or counter-clockwise rotating one.

There are some differences between polarization phenomena viewed by optical specialists and antenna designers. In the theory of antennas, there exists horizontal and vertical polarization, meaning in fact a linear polarization with the end of the electric field in a horizontal plane (horizontal polarization) and along the vertical one (vertical polarization). There is also a difference between the clockwise and counter-clockwise rotating polarizations [3].

In the literature there are two main methods to characterize propagation: the so-called *Forward Scattering*

Alignment (FSA) and the so-called *Backward Scattering Alignment* (BSA). FSA is used to determine a change of polarization of a wave propagating to and from an observed target through the surrounding medium. The polarimetric scattering properties of the medium are described in this case by its Jones or Muller matrices. BSA is especially useful when analyzing transmission through a scattering object between antennas of various polarizations. Sinclair and Kennaugh matrices are in use to describe polarimetric properties in this case [5]. Polarization has a dual aspect: on the one hand, it describes the behavior of a complex harmonic vector field, on the other hand it serves as a way to describe scattering properties of different propagation media.

In order to characterize this aspect of the radio wave propagation, the role of precise wave propagation models is particularly important as through measurements is not possible to accurately measure the polarization because we are limited to certain types of antennas. In other words the radiation pattern of the antenna will influence our measurements and will not allow us to gain precise information regarding it. For the model that describes the polarization, a proper definition for the case of complex harmonic vector fields and an adequate set of graphical and numerical representation is needed. Our contribution addresses this problem.

3. The Stokes-parameters and the Poincaré-sphere

We can introduce the Stokes-parameters and the Poincaré-sphere [2,3] to characterize the polarization of an arbitrary plane wave. The Stokes-parameters are defined as follows:

$$s_0 = a_1^2 + a_2^2$$
$$s_1 = a_1^2 - a_2^2$$
$$s_2 = 2a_1a_2\cos\delta$$
$$s_3 = 2a_1a_2\sin\delta$$

where a_1 and a_2 represent the two perpendicular components of the field strength, one in the plane of the receiver and the other in the plane vertical to the receiver's plane and to the propagation direction. And only three of those four are independent because:

$$s_0^2 = s_1^2 + s_2^2 + s_3^2 .$$

The Stokes-parameters generally are defined by four element vectors: $S = [s_0 \;\; s_1 \;\; s_2 \;\; s_3]$.

For a horizontally polarized linear plane wave we have: S=[1100]
For a linear plane wave with circular polarization (45°): S=[1010]
For a linear plane wave with clockwise circular polarization: S=[1001]
For a linear plane wave with counter-clockwise circular polarization: S=[100-1]
For a non-polarized wave: S=[1000]

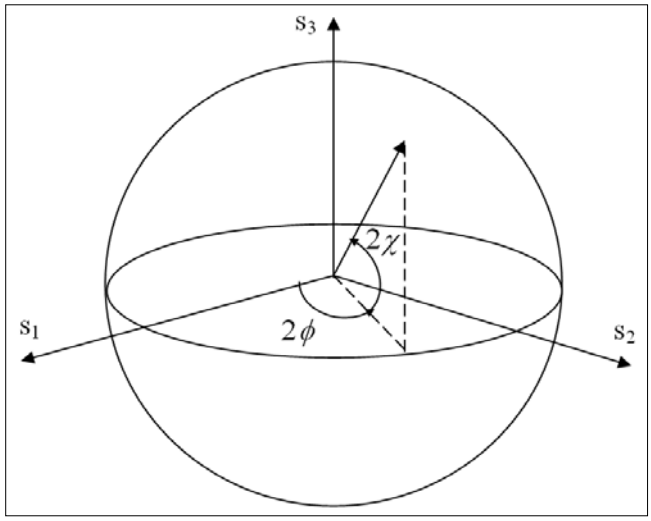


Figure 1.
Characterization of the Stokes-parameters and the Poincaré-sphere

We can define the triplet $\sqrt{s_0}, \chi, \phi$ called ellipsometric parameters. We can establish the following relationships between them and the Stokes-parameters:

$$s_1 = s_0\cos 2\chi\cos 2\phi$$
$$s_2 = s_0\cos 2\chi\sin 2\phi$$
$$s_1 = s_0\sin 2\chi$$

4. Polarization ellipse distribution and the direction of the major semiaxis

The Poincaré-sphere gives a very good visual representation of the various polarization states of a wave: one point on the sphere corresponds to every possible state of a plane monochromatic wave of a given intensity s_0 and vice versa. However, in a complex indoor environment we do not have constant field intensity, neither monochromatic waves, but there are multipath components. So the Poincaré-sphere is not adequate for a proper description of such a complex harmonic vector field.

We use another approach [3] for a more detailed description, assuming that the electromagnetic field provided by a harmonic source that penetrates through a building, which suffers multiple reflections, transmissions and diffraction, remains harmonic. Therefore in every point of the space we have a harmonic vector, **V**, a three-dimensional function of the field intensity in that point:

$$V_x(r,t) = a_x(r)\cos(\omega t - \varphi_x(r))$$
$$V_y(r,t) = a_y(r)\cos(\omega t - \varphi_y(r))$$
$$V_z(r,t) = a_z(r)\cos(\omega t - \varphi_z(r))$$

where

for example $\varphi_x(r) = k \cdot r - \delta_x$ and so on meaning the initial phases of the scalar components of the vector field.

We can introduce the following vectors: $p(r)$ and $q(r)$, which depend on the position vector, as follows:

$$p_x(r) = a_x(r)\cos\varphi_x(r), p_y(r) = a_y(r)\cos\varphi_y(r)$$
$$p_z(r) = a_z(r)\cos\varphi_z(r), q_x(r) = a_x(r)\sin\varphi_x(r)$$
$$q_x(r) = a_y(r)\sin\varphi_y(r), q_x(r) = a_z(r)\sin\varphi_z(r)$$

Then we have:

$$V(r,t) = p(r)\cos\omega t + q(r)\sin\omega t$$

Between these two vectors there exists a relationship: they are conjugate semi-axes of the polarization ellipse. If we choose two other vectors, *s*(*r*) and *u*(*r*) with the following relations:

$$s = p\cos\alpha + q\sin\alpha$$
$$u = p\sin\alpha + q\cos\alpha$$

and we choose α such as *s* and *u* become perpendicular,

$$tg2\alpha = \frac{2pq}{p^2 - q^2}$$

then *s* and *u* will be the major and minor semi axes of the polarization ellipse. To see this, we can express *V* in terms of *s* and *u* as:

$$V = s\cos(\omega t - \alpha) + u\sin(\omega t - \alpha)$$

Now that we have *s* and *u* perpendicular one to the other, we can choose a new coordinate system with the *x'* and *y'* axes along *s* and *u*. In this new coordinate system the components of vector field *V* will be the following:

$$V_{x'} = |s|\cos(\omega t - \alpha),\ V_{x'} = |u|\sin(\omega t - \alpha),\ V_{z'} = 0$$

Therefore we can conclude that:

$$\frac{V_{x'}^2}{s^2} + \frac{V_{y'}^2}{u^2} = 1$$

so $s=|s|$ and $u=|u|$ will be the major and minor semi-axes of an ellipse that describes the vector field in the rectangular coordinates defined by these vectors, which means that in fact it is the polarization ellipse itself.

So in the most general case in every point of the considered space the polarization is elliptic, but the plane of the ellipse, the direction of the greatest diameter within the plane and the eccentricity of the ellipse varies. That can be easily seen: as α varies, the direction of axes *x'* and *y'* also varies and so does the direction of the ellipse's support plane.

In conclusion we propose the following graphical representations for the polarization analysis:

– the distribution of the ellipse eccentricities, giving a description of how diffuse the field gets inside the building due to the obstacles and different propagation mechanisms;
– the two-dimensional histogram of the major semi-axis of the ellipses, varying φ and θ, it could give information related to the needed orientation of the receiver's antenna.

5. Simulation results and analysis

Using our ray-tracing tool [7,8], we simulated the radio waves generated by an external transmitter (LEO satellite) inside a building. Multiple reflections and transmission through objects were taken into account; diffraction is not included in our model so far. The transmitter generates waves with different polarizations, at 2.4 GHz (downlink band assigned for LEOS by WARC'92): circular (clockwise and counterclockwise) and linear.

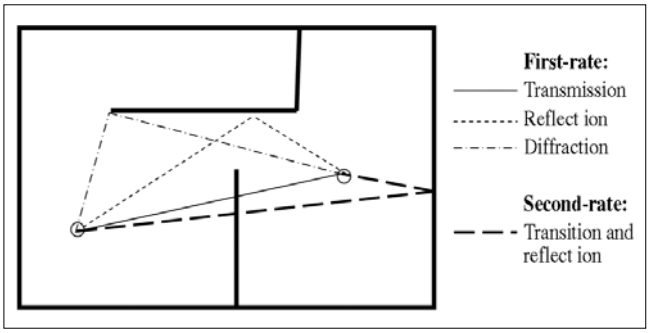


Figure 2.
Some possible first- and second-rate component of the Ray Launching

The ray launching wave propagation models are based on geometrical optics instead of the full space modeling. The propagated waves are divided into finite space angles and these components are treated independently. The model provides a complete result from point to point by the independent space components and the phenomena on the different surfaces (reflection, transmission, diffraction).

In practice the method of the ray-launching is extended for third-rate arbitrary propagation mechanism combination (in our simulation we took into account seventh-rate combination) or we follow the wave while the field strength of the followed wave decreases under a definite level. Our ray launcher is based on the ray launching concept; it has at its origin the Luneberg-Klein [2] series expansion, a high frequency approximation called geometrical optics.

	Ray Launching
Characteristics	• Frequency domain method • Narrow-band sinusoidal excitation
Advantage	• Easy partitioning
Disadvantage	• Programming is complicated • Significant ray divergence for complex geometries

Table 1. Characteristics of ray launching

The ray launching in our case has been applied in inverse direction: we launch rays from the receiver locations under 8 different angles,in which directions the rays propagating from the satellite could arrive. We follow the rays until a given number of intersection points is reached, which is 7 in our simulation. We took into account only those rays that propagate in the direction of the satellite position, so basically a plane wave, because the real wave, propagating from the satellite towards the building, can be considered as plane due to the great distance between the satellite and the building.

We present simulation results for these polarizations, in the near field of the building side, at different distances from the windows, for the case of satellites at various elevation angles.

The figures present simulated results of the polarization characteristics, for two regions of the building: one inside the rooms that were directly irradiated through windows and the next one for in-building regions not directly irradiated, inside the building, for low elevation angles and for clockwise, counterclockwise and linear polarization, respectively, of the wave source.

At high elevation angles (45° and above) the principal penetration mechanism is the diffraction through window frames, as transmitted rays would rapidly bounce between the ceiling and the floor, rapidly attenuating, therefore rather few rays arrive into remote regions, so generally valid conclusions cannot be drawn for these angles so far.

Low and medium elevation angles at which the penetration through windows is the principal mechanism, can also be taken into account. The results show that the wave, suffering specular multiple reflections and transmissions, changes its polarization state, in other words the polarization of the incident plane wave will be preserved only in the near-window region; otherwise, in all three examined cases, i.e. for linear, clockwise and counterclockwise polarizations of the incident wave, in distant regions the polarization ellipses become almost evenly distributed, so it seems that there is no direct connection between the polarization state of the incident waves and the polarization of ellipses in different points of the building. As for major semi-axis orientation of the polarization ellipse, it varies rather deeply. However, the distribution curves can be observed to be centered at the angles corresponding to those under which the satellite is seen, or those under which once or more than once the reflected waves reaching the satellite, generate from the source.

6. Simulation results for satellite-MIMO channel

In our work, we analyzed the MIMO – satellite channel in scatterer and non-scatterer environment. The transmitter was a three – element dipole antenna system and

Figure 3.
25°, 5°, near zone, clockwise, 50% of points clockwise

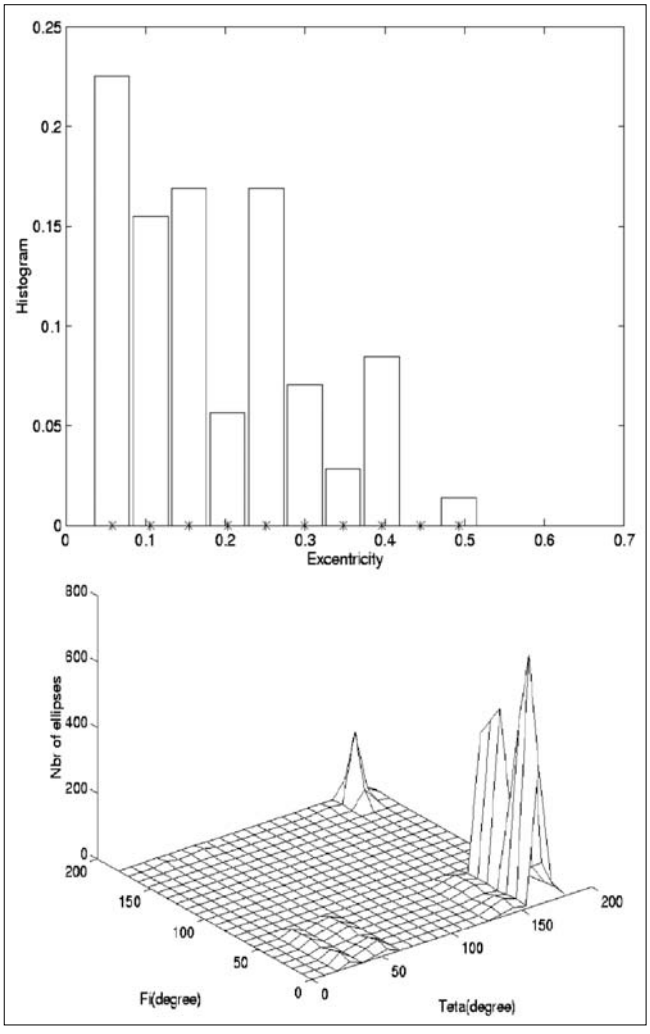
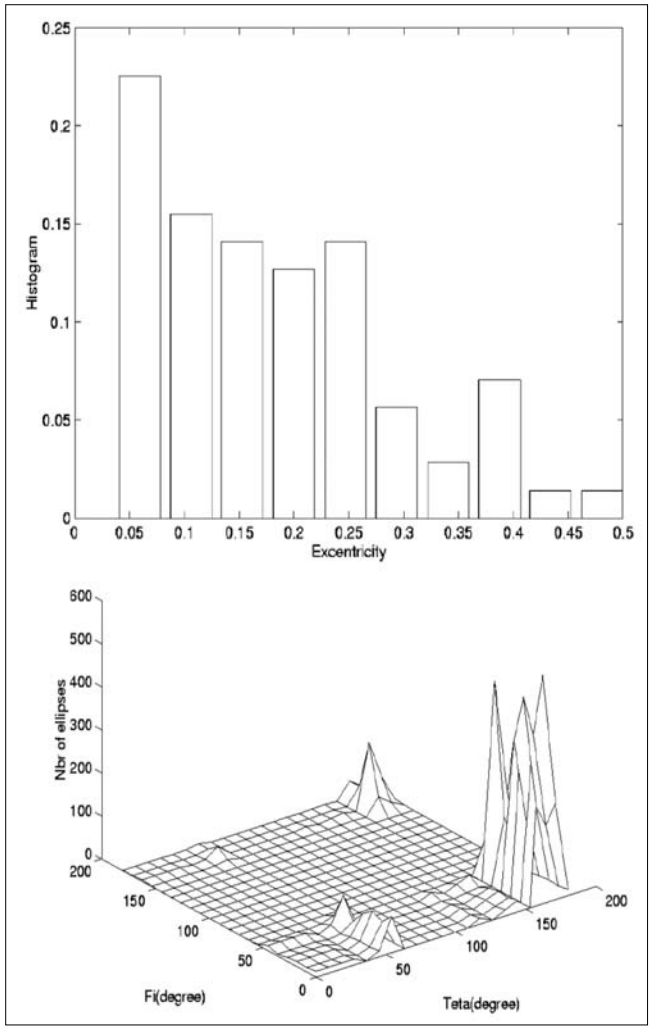


Figure 4.
25°, 5°, near zone, counterclockwise, 48.5% of points clockwise



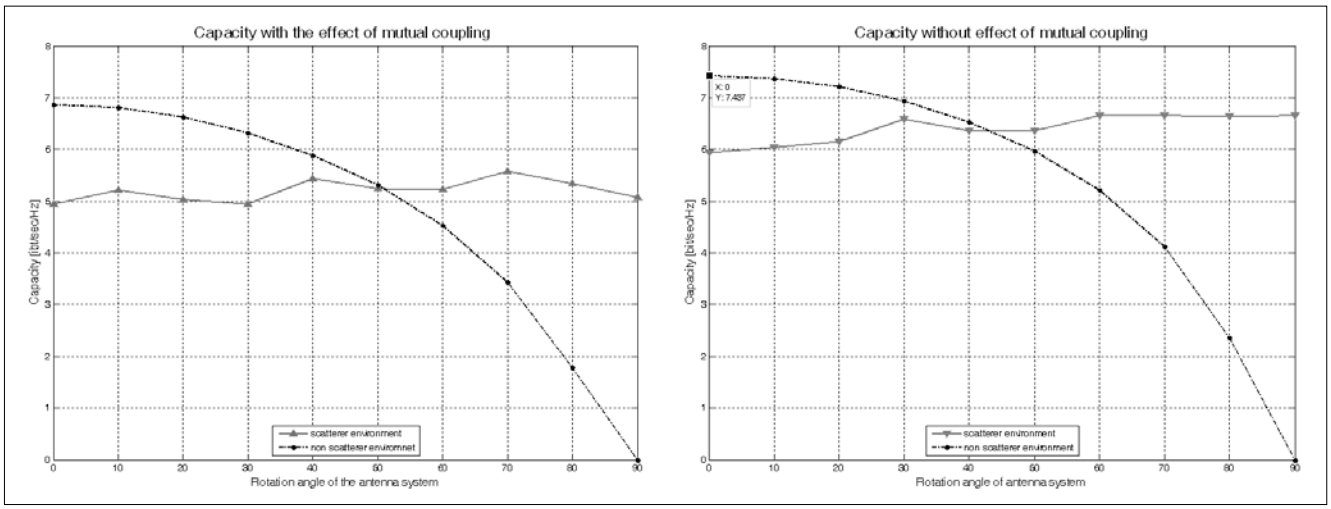


Figure 7. The capacity for scatterer and non-scatterer environment

the receiver was the satellite. This is why the simulated system was a SIMO structure. In the course of the simulation the receiver antennas were rotated. At first the tree antennas are in line with the axel-Z (rotation angle 0°). Then we opened the antennas like an umbrella the end position was the plane X–Y (rotation angle 90°). There was 120° between the projections of the antennas on the X–Y plane.

The scattered environment is an inside place where object can reflect and scatter the waves which came

Figure 5.
25°, 5°, far zone, counterclockwise,
45.1% of points clockwise

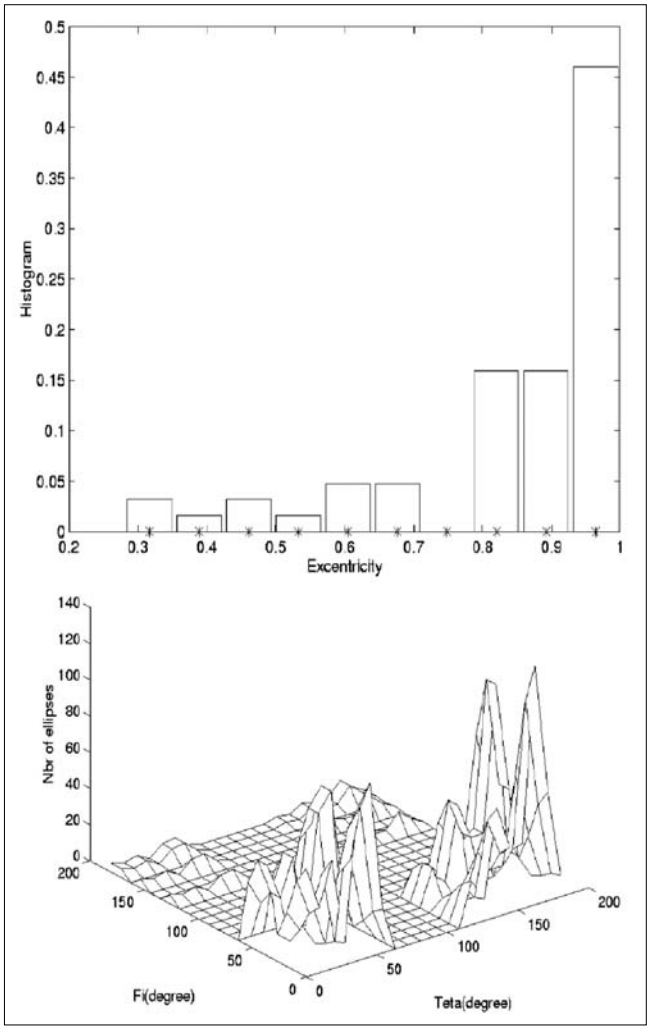
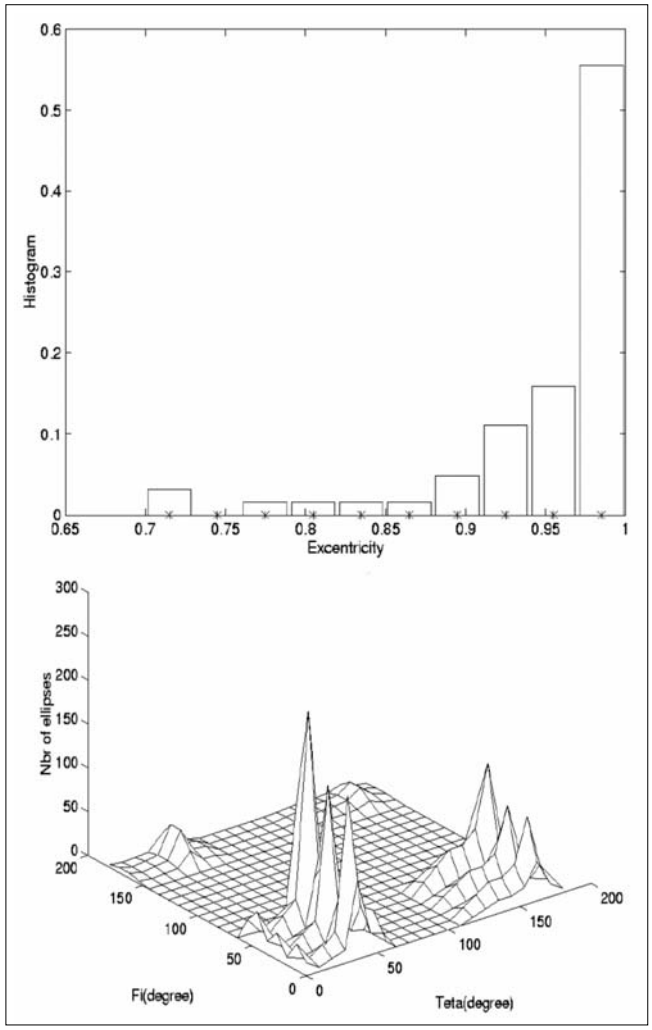


Figure 6.
25°, 5°, far zone, linear,
54.8% of points clockwise



from the satellite. The non-scattered environment is a so special reference inside place in where there is not scattered object. Consequently there is not reflection or scattering between the transmitter and the receiver unit.

Figure 7 shows the result of the simulation. The broken line denotes the channel capacity in a non-scattered environment, and the continuous line is the channel capacity for standard environment. Without reference to the result of the mutual coupling it is evident that the channel capacity is about moderate in scatterer environment and dynamically fluctuates in non-scatterer environment. Consequently the scatterer environment causes increase in the channel capacity by n×m channel.

7. Conclusions

In our work we presented a method for describing the polarization state of a complex harmonic electromagnetic field inside a building. The presented graphical representations can be used to get an insight into the complex polarization phenomena of the radio waves in the case of multipath propagation environments, with application for the satellite-to-indoor radio propagation channel.

A generic conclusion is that, for an office type building, the polarization state of an incident plane wave does not have a major effect on the complex indoor harmonic field. Its polarization is not preserved in the far-window regions. In other words a circularly or linearly polarized plane wave generates a circularly or linearly polarized complex harmonic field inside the building only in a rather close proximity of the penetration regions, otherwise the field is diffuse.

Acknowledgement

We would like to thank MIK (Mobil Innovation Center) for its support.

Authors

LÓRÁNT FARKAS received his MSc degree from the Technical University of Timisoara, Romania in 1996, and his MBA degree from the Budapest University of Technology and Economics in 2005. Currently he is a Senior Researcher at Nokia Siemens Networks Hungary. His interests include streaming solutions, distributed multimedia systems and web technologies.

LAJOS NAGY has finished his studies at the Budapest University of Technology and Economics, specialization telecommunications in 1986 and his post-graduate engineering studies in 1988, obtaining a diploma with distinction. He has a Dr. Univ. degree (1990) and a Ph.D. degree (1995). At present he is Assoc. Professor and Head of the Department of Broadband Infocommunications and Electromagnetic Theory at the Budapest University of Technology and Economics. His research interests include applied electrodynamics, mainly antenna design, optimization and radio frequency propagation models. Dr. Nagy is Secretary of the Hungarian National Committee of URSI and the Hungarian delegate to the Section C of URSI. He is leading the Hungarian research teams in COST 248 and ACE2 EU projects. He has published over 100 papers.

ANDREA FARKASVÖLGYI graduated from the Budapest University of Technology and Economics, Faculty of Electrical Engineering, in 2002. She finished the Ph. D. school at the Department of Broadband Infocommunications and Electromagnetic Theory in 2006. Her research interests include satellite systems and MIMO antenna systems.

References

[1] IEEE Standard 145-1983, IEEE Standard Definitions of Terms for Antennas.
[2] C. E. Balanis, "Advanced Engineering, Electromagnetics", Wiley & Sons, 1989., pp.154–173; 748–760.
[3] M .Born, E. Wolf, "Principles of Optics", Pergamon Press, 1975., pp.28–36.
[4] G. R. Hoeft, "Ground Penetrating Radar", (1998), http://www.g-p-r.com
[5] Z. H. Czyz, "Polarimetric Bistatic Scattering Transformations as seen from Two Different Points of View: Optical (Propagation) and Radar (Transmission) – The Poincaré Sphere Analysis", U.R.S.I. General Assembly, 1999., p.359.
[6] D. J. de Smet, "A Closer Look at Nulling Ellipsometry", 1995., http://www.tusc.net/~ddesmet
[7] Z. Sándor, L. Nagy, Z. Szabó, T. Csaba, "Propagation Modeling", Microwave and Optics Conference, MIOP'97, Sindelfingen, Germany, 1997., pp.213–215.
[8] Z. Sándor, L. Nagy, Z. Szabó, T. Csaba: "3D Ray-Launching and Moment Method for indoor Radio Propagation Purposes", The 8 International Symposium on Personal Indoor and Mobile Radio Communications, PIMRC'97, Helsinki, Finland, 1997., Vol. I., pp.130–134.
[9] Adrián K. Fung, "Microwave Scattering and Emission Models and their Applications", Artech House, 1994., pp.14–26.
[10] B.G. Molnár, I. Frigyes et al, "Characterization of the Satellite-to-Indoor Channel based on Narrow-Band Scalar Measurements", PIMRC'97, Helsinki, Finland, Vol. 3., pp.1015–1018.
[11] Raymond L. Pickholz, "Communications by means of Low Earth Orbiting Satellites", Modern Radio Science, Oxford University Press, 1996. pp.133–151.

Test of the SAS2
ULF-VLF electromagnetic wave analyzer in space environment – on board of the Compass-2 satellite

CSABA FERENCZ, JÁNOS LICHTENBERGER, ORSOLYA E. FERENCZ, DÁNIEL HAMAR
Space Research Group, Lóránd Eötvös University of Sciences, Hungary; spacerg@sas.elte.hu

LÁSZLÓ BODNÁR[1], PÉTER STEINBACH[2], VALERY KOREPANOV[3]
[1]*BL-Electronics, Solymár, Hungary; bodnarl@bl-elelctronics.hu*
[2]*MTA-ELTE Research Group for Geology, Geophysics and Space Sciences; spacerg@sas.elte.hu*
[3]*Lviv Centre of Institute of Space Sciences, Lviv, Ukraine; vakor@isr.lviv.ua*

GALINA MIKHAJLOVA, YURI MIKHAJLOV, VLADIMIR D. KUZNETSOV
IZMIRAN, Troitsk, Moscow Region, Russia; yumikh@izmiran.ru

Keywords: VLF, whistler, electromagnetic wave propagation, electromagnetic monitoring, space weather, SAS instrument

The SAS2 measuring system, an advanced version of the SAS electromagnetic wave analyzer and sampler, successfully operated on board of the Compass-2 satellite launched in 2006. The main mission of this satellite was to be a technical test of a satellite series in space environment. These satellites are intended to research space weather, further they will observe possible precursory events of earthquakes. The measuring systems and instruments worked well. We found some interesting phenomena in the observed ULF-VLF electromagnetic database detected by SAS-K2: whistler-doublets (observed earlier in 1989, by the first version of SAS on board of IK-24), "spiky" whistlers (SpW, we identified this signal type first from database of Demeter satellite, and we presented the theoretical solution of these signals). Further signals, propagated in ducted whistler mode between two inhomogeneous surfaces (onion-skin) in the plasma, were successfully identified first time during this mission. These signals presumably propagated in higher mode (third order) in the magnetosphere. We delivered the theoretical solution of this phenomenon for UWB (ultra wide band) signals too.

1. Introduction

The monitoring of the electromagnetic environment of planets (primarily the Earth) essentially starts nowadays, however some experiments were conducted in the past. On the one hand, the required measuring and data-handling techniques have become available in the recent years; on the other hand, the needed wave propagation theory has been born only by now, as many theoretical breakthroughs were necessary in the case of these problems.

One of the early experiments was the IK-24 ("Active") satellite launched in 1989 [8], which carried onboard the Signal Analyzer and Sampler (SAS) instrument developed by us. By now it became obvious that the investigation, observation and continuous monitoring of the electromagnetic surroundings of planets and especially the Earth are undoubtedly necessary in order to understand and to model the processes of the planet. For this task we developed the advanced versions of SAS measuring system, the SAS2 and SAS3 (SAS2 for satellites and interplanetary space probes, while SAS3 for satellites and for board of the ISS, in order to detect very high resolution of waveforms). By the application of these data it is possible to obtain a more precise description of the terrestrial processes, to have a better understanding of the solar-terrestrial connections and the space weather, further to construct an adequate database for prediction of terrestrial seismic activities and earthquakes.

First of all, for the better understanding and application of signals registered onboard, and for deducing right conclusions from their shapes (i.e. evacuation of a city because of a likely earthquake) we have to describe the generation and propagation of these electromagnetic signals. As an example, let us consider one of the important problems, the ULF-VLF signals excited by lighting discharges (the whistlers). A highly accepted theory for the generation and propagation of these signals (e.g. [7]) says that these signals can only and exclusively propagate from the location of the exciting lightning to the conjugate point (the location of detection) of the geomagnetic field line on the Earth's surface in waveguides formed by ("spaghetti-like") plasma inhomogeneities (ducts). Another hypothesis assumes that these signals can get to the conjugate point without ducts. A third opinion considers onion-skin type inhomogeneities in the high atmosphere as waveguides.

The answer of this question is important for the right interpretation of the observed phenomena. In order to answer this, we have to know exactly the method of the propagation and the shape of the UWB signal [1] in this special environment, along with the different waveguide models. By the application of a sufficiently rigorous theoretical model and by simultaneous, continuous and automatic terrestrial and onboard measurements, the answer can be found [2,5]. Besides the continuous and effective developing of our theoretical models and solution methods, the Demeter mission and the advan-

ced SAS2 with Compass-2 were essential steps forward in this way. Furthermore, we have started to install a global terrestrial measuring system (VR-1 and VR2), compatible with SAS2. This measuring network has been successfully operating at four points in the Carpathian basin, at two points at the Antarctica, two points in South Africa, in New Zeeland, and other places (in the near future e.g. two points in Finland). This terrestrial network works continuously, automatically recognizes, evaluates and classifies the whistlers, and since the first half of this year automatically delivers the plasma parameters, which was unavailable earlier. This is the AWDA system [11]. For this development it was necessary to work out the most accurate, new UWB wave propagation models, because the former theoretical approaches were not applicable for this task.

First we briefly describe the goals of Compass mission and the advanced SAS2, and then we outline the exact UWB description of the guided signal propagation in waveguides filled by magnetised plasmas and the clear evidence of the guided-mode propagation in the detected data of SAS2-K2 worked on board of Compass-2.

2. Advanced Signal Analyzer and Sampler in the Compass mission

The main goal of the Compass mission was to test and verify the measuring ideas and complete detector and instrumentation system of planned missions with the same scientific goals following the Compass. The general scientific goals of the complete mission are the research of the electromagnetic activity concerning the seismic events, especially the electromagnetic precursors of the earthquakes, the detailed investigation and monitoring of the electromagnetic environment of the Earth in the ULF-VLF bands, inside this area the detailed characteristics of the lightning activity and whistler propagation, and space weather relations.

The Compass-1 was launched on 10th December 2001, however, after the successful launch the satellite failed on orbit. The Compass-2 was launched on 26th May 2006 and reached the planned low circular orbit (inclination 79°, orbit height ~400 km). See the satellite in *Fig. 1* and the general sketch of the satellite in *Fig. 2.*

After the start of the operation serious problems appeared in the power system of the satellite, which was partly solved by November 2006. From that time during the active life of the Compass-2 the scientific sensors and measuring systems worked perfectly.

One of the scientific measuring equipments is the advanced Signal Analyzer and Sampler (SAS2-K2, developed and produced by the Space Research Group of the Eötvös Loránd University and the BL Electronics) using Ukrainian sensors, two spherical electric sensors producing one (differential) electric signal appearing between the two spheres, and one electric search coil sensor producing one magnetic signal in the ULF-VLF bands. One of the sensors (ES) has two conductor balls positioned >1.5 m far from each other, the measured signal is the electric potential difference between the balls. The other sensor (MS) is a search coil receiving ULF-VLF magnetic signals parallel with its axis. The SAS2-K2 and the sensors worked perfectly with high symmetry between the electric and magnetic channels, with high sensitivity and low noise according to the original specifications.

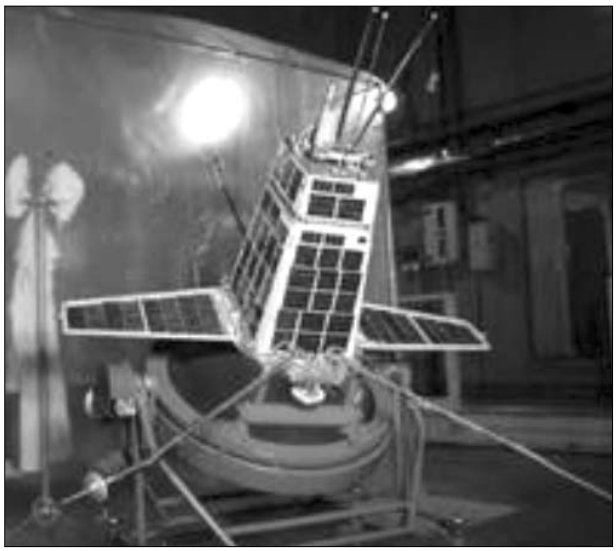


Figure 1.
The Compass-2 during the final phase of the ground tests

Figure 2.
The sketch of the Compass-2 satellite. The sensors of the SAS2-K2 are the two ES electric and the MS magnetic Ukrainian sensors.

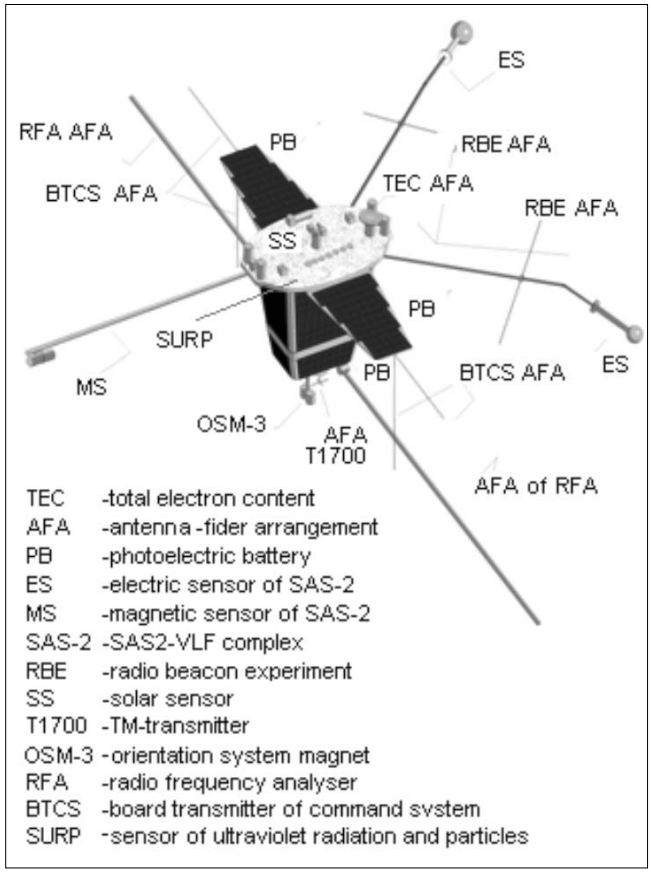


Figure 3.
The SAS2-K1 equipment for the Compass-1 satellite

This fact means a qualitative step forward in comparison with Demeter satellite, because the sensitivity of the magnetic channel on the Demeter is much lower than the electric one. Both the electric and the magnetic components of Compass-2 have been registered and evaluated with the same quality. The SAS2-K1 for Compass-1 and the SAS2-K2 for Compass-2 can be seen in *Fig. 3* and *Fig. 4.* Inside the box of SAS2-K2 two complete and identical SAS2 were integrated to increase the reliability of the whole mission, as it is possible to see in the block-diagram of the system *(Fig. 5)*.

The original SAS2-K1 launched on board of Compass-1 had only one, two-channel SAS2 module (see in Fig. 3). The SAS2-K2 contained two, identical, two-channel SAS2, one active and one cold backup, in order to increase the reliability of the whole mission. This can be well seen in the block diagram (Fig. 5).

The main characteristics of the SAS2 wave analyzer are the following:

Frequency range
(search coil and the electric spheres, 1 pair):
1 Hz – 20 kHz
Search coil transfer function:
1 Hz – 1 kHz linear, 1 kHz – 20 kHz flat
Electric sensor transfer function:
1 Hz – 20 kHz nearly flat
Noise bands:
Magnetic sensor: 10 Hz – 2 pT / Hz1/2
100 Hz – 0.2 pT / Hz1/2
1 kHz – 0.03 pT / Hz1/2
10 kHz – 0.05 pT / Hz1/2
Electric sensor (1 pair):
10 Hz – 40 nV / Hz1/2
10 kHz – 20 nV / Hz1/2

The mass of the SAS2-K2 electronic unit (Fig. 3b) is 470 g. The total mass of the system with sensors etc. is 1260 g. The size of the electronic unit is 150x70x110 mm. The power consumption is ≤ 3 W, including the sensors, however, one electronic unit (the SAS2-A or the SAS2-B inside the SAS2-K2 box) is a cold backup.

During the operation of the system it is possible to select, by ground commands, that the A or B unit is active inside the SAS2, to select the input gain of the input analogue (ULF-VLF) amplifiers and to select the main operation mode and the parameters of the selected operation mode including the sampling rate of the incoming (registered) signals. (The maximum sampling speed is 43.2 kHz for each channel and the gain of input amplifiers is changeable from -18 dB to +20 dB in three steps.

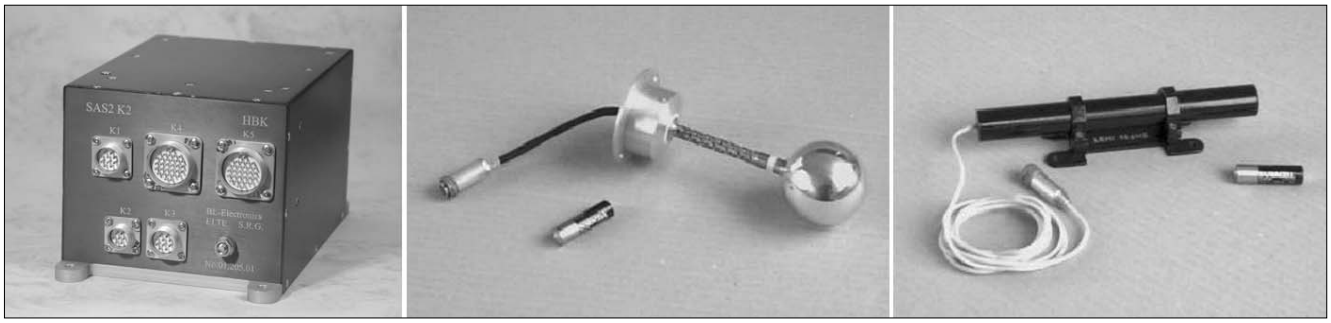
The SAS2 has three different memory modules beside the internal memory of the DSP. The first is the boot memory (128 kBytes EPROM) from which the operating program is loaded after the switch-in or power-on reset. The second is a 4 MBytes SRAM used for measurement data storage as circular buffer and telemetry buffer. The third is a 64 kBytes EEPROM to set and store the actual parameters and reference spectrum of the measuring (operation) modes.

Using this hardware and software possibilities, the SAS2 main operation modes are the following:

a) Monitoring of the average electromagnetic noise spectra of the two channels. The averaging time can be set by commands, by parameters from 1 sec high speed monitoring to 10 minutes long time averaging. This operation mode is running simultaneous with the *b)* or *c)* operation modes.

b) Triggered event detection: After the processing of the signals incoming simultaneously in the two channels, the processed signal (spectrum) of one selected channel is compared to a stored reference signal (spectrum) which is selected or modified by ground commands according to the actual onboard EMC etc. If the recorded spectrum is higher than the reference one on one or more spectral lines (see more about these criteria in [11]), it is accepted as a real event and valid trigger. Then the system reads the recorded signal values stored in circular buffer memory before a predefined time of the trigger signal and following the trigger signal to another predefined time (e. g. 0.5 sec before the trigger and e. g.

Figure 4.
The SAS2-K2 equipment for the Compass-2 satellite and the electric and magnetic sensors of the SAS2-K2, respectively



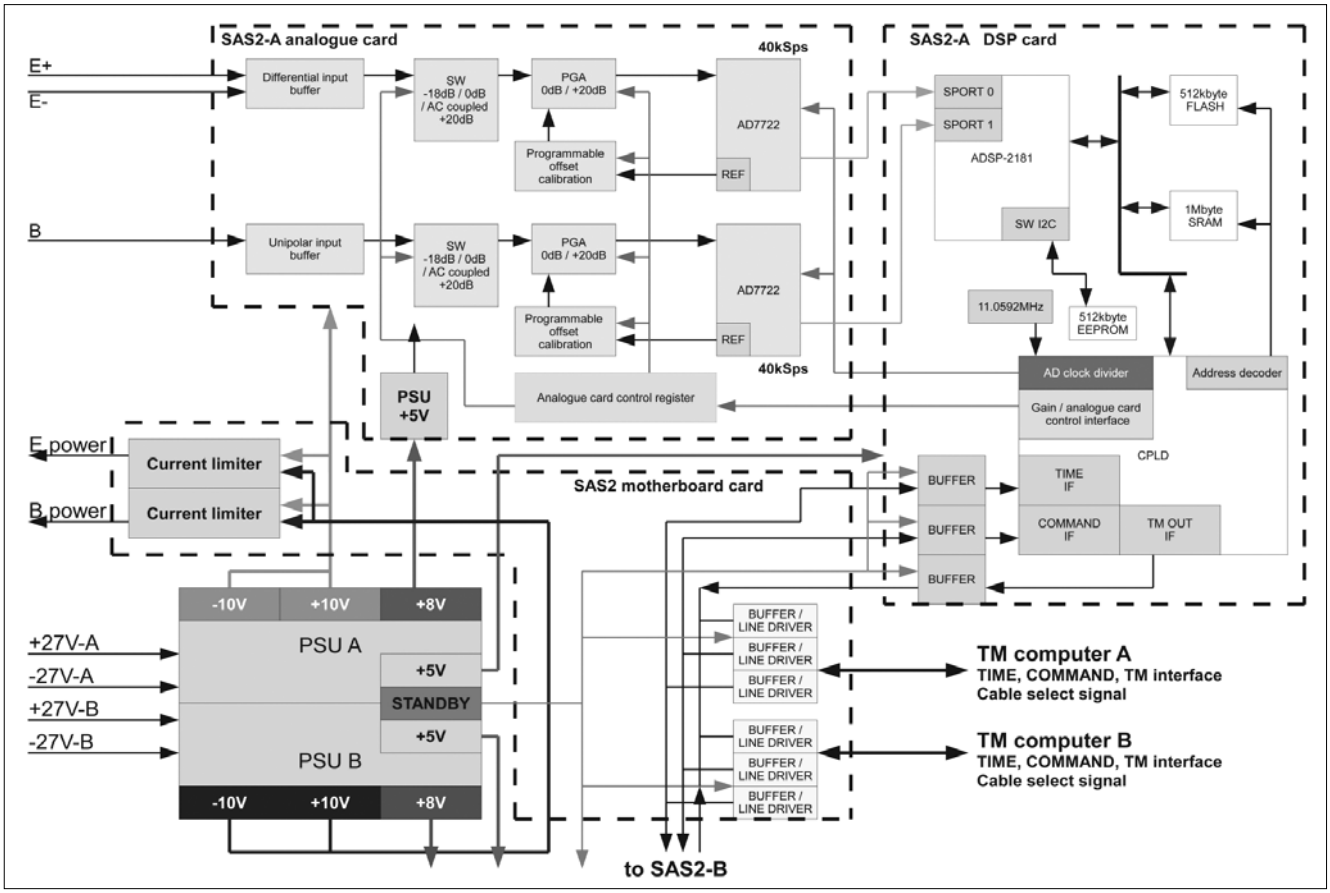


Figure 5. The block diagram of the SAS2-K2. The SAS2-A and SAS2-B units are identical

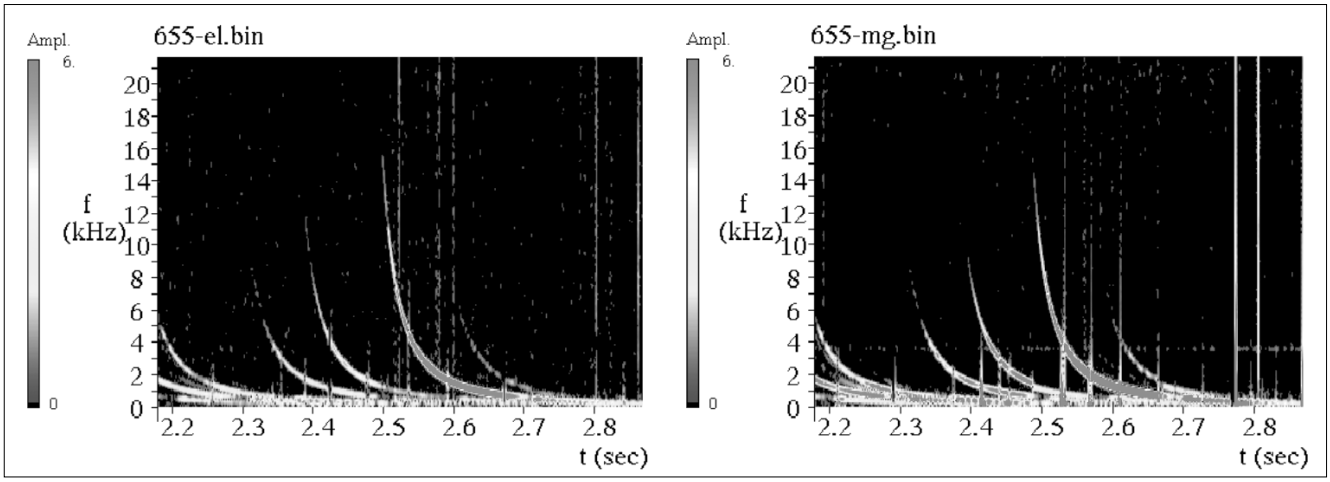
1 sec or 2 sec after the trigger). If the registration of the detected electric and magnetic signals during this time period is finished, the whole record is rewritten into the telemetry buffer.

c) Periodical, time controlled data collection: In this mode the SAS2 detects and stores the incoming data of the two channels with predefined sampling rate (normal "burst" mode) without triggered event detection using a command controlled (predefined) registration time schedule list.

The SAS2 system worked perfectly, and the technological test of the planned complete measuring system was successful, too. The SAS2 registered ULF-VLF signals with the planned specification, with low system noise and the electric and magnetic channels have the symmetry in the signal levels of the incoming electromagnetic events and noise, which is important in the electromagnetic research (see e. g. *Fig. 6*).

In Fig. 6, the first measured signal, a very intensive whistler group can be seen, with both the electric and

Figure 6.
One of the first data set measured by the SAS2-K2 on board of Compass-2, 29th Nov. 2006. UT 5.00.00 above Indonesia. We see here a very strong whistler group, i. e. the FFT pattern of the electric and magnetic components of these whistlers, respectively. The symmetry of the sensitivity of the detector channels is evident.



the magnetic components. In the followings we show briefly three types of detected signals and their interpretations.

3. Examples from the electromagnetic events registered by SAS2

In the followings three examples are displayed from the SAS2 registrations. The first and the second data correspond to our previous theoretical results, the third one shows a new and unknown phenomenon.

3.1 Whistler-doublets

The first detected whistler-doublets appeared during the first SAS experiment, on board of the IK-24 ("Active") satellite – see *Fig. 7*, and more in [8]. The origin of this phenomenon was supposed to be a simple reflection arriving upward below the satellite – this hypothesis seemed to be supported by the time shifting of the two traces and the position of the satellite –, or to be a traveling time difference caused by the presence of two, neighboring plasma inhomogeneities (ducts). The time difference between the two traces in a whistler was 70 < 80 ms in the case of this data, which could make both explanations to be likely. It was also probable, on the basis of the very similar fine structures of the traces that the source (the generating lightning discharge) was identical for both signals, and the propagation paths were closed. But the verification of this interpretation by real wave propagation model calculations has been missing up to now, as the first really exact UWB propagation models were developed later.

The SAS2-K2 has detected whistler doublets, too (see in *Fig. 8*). The time difference in this case is 60 < 70 ms. However in this measured data there were no more doublets in 1-2 seconds, but similar doublets were registered in the same measuring record few minutes earlier. It means that whistler doublets can occur occasionally, but not systematically. The phenomenon is not accidental. Nevertheless, our former interpretation needs fundamental revision, because the orbit of the Compass-2 satellite is much lower than in the case of IK-24, the height of the orbit is ~400 km. This height is too low for such a big running time delay of a signal reflected backward to the satellite near the surface.

The Compass-2 orbits in the ionosphere, just like that of the Demeter satellite, are not too much above it. So, the way of generation by reflection can be excluded. But the other generation theory regarding the two, near, narrow waveguides with small diameters can also be excluded on the basis of the exact UWB solution. We have successfully deduced the UWB solution for waveguides filled by anisotropic, magnetized plasma and the registered doublets show no typical signs of guided propagation (the guided signals have typical forms, see Section 3.3). Thus it became necessary to reexamine this problem, and to find a new, consistent explanation of their generation (as the occurrence of the exciting lightning is possibly in connection with global climate processes and their changing).

3.2 Spiky Whistlers – SpW

We briefly reported this type of whistlers in our paper on registrations of Demeter satellite [2]. The exact identification and wave-theoretical description of this phenomenon, together with the whole UWB solution and model can be found in [3,5].

Here we summarize the essence of this theory. The lightning discharge excites dominantly a vertical current, so-called cloud-to-ground directed lightning. The excited electromagnetic signal starts to propagate in the Earth-ionosphere waveguide, and after traveling a longer distance in this waveguide the signal can connect out toward the higher atmosphere. Through the magnetized ionosphere plasma, the signal can reach the satellite. Because of the fact of guided propagation in the Earth-ionosphere waveguide, guided modes of different orders can appear in the spectrum, with different, harmonic wavelength limits depending on the distance between the Earth's surface and the bottom of the D-layer of the

Figure 7.
Whistler-doublets detected by first SAS experiment on board of IK-24, 14th December, 1990.

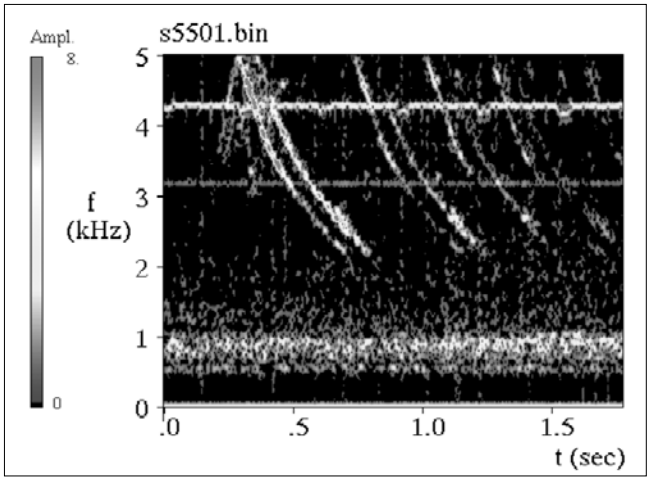
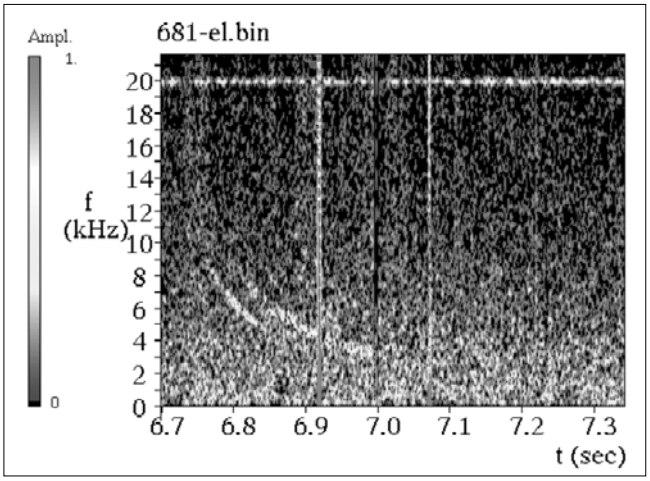


Figure 8.
Whistler doublet measured by SAS2 on board of Compass-2, 27th of January, 2007.



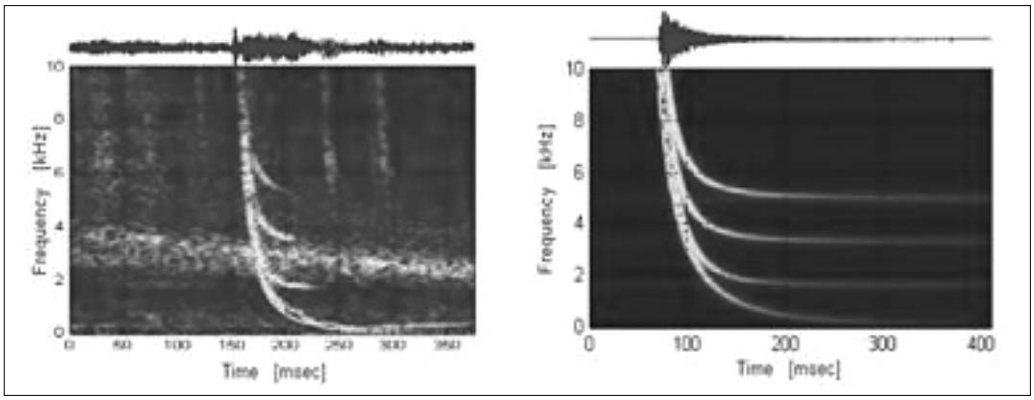


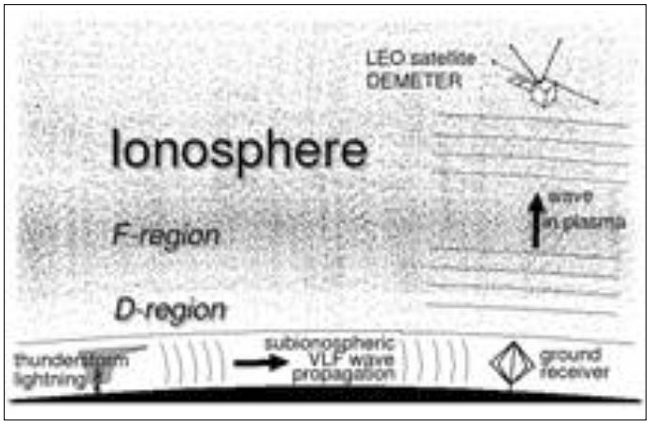
Figure 10.
FFT spectrograms and signal forms of a measured (Demeter satellite, 6th of November, 2004) and a calculated UWB signals [5]

ionosphere. Thus the spectrogram of the signal (e.g. the FFT pattern) will be "spiky". This signal suffers dispersion during propagation toward the satellite, but preserves its original spiky structure. The sketch of the propagation can be seen in *Fig. 9.*, the signal calculated by our UWB model in comparison with data detected by Demeter satellite can be seen in *Fig. 10.*

We have also detected SpW-s on board of Compass-2, see *Fig. 11*. In this signal group, the second, third, sixth, seventh and ninth order modes can be identified and classified. From these modes, the third and ninth order modes are very sharp, the others are recognizable with small intensity, and no more modes propagated toward the satellite in this case.

We successfully verified the fact, that the Spw-s, which we identified first from registrations of the Demeter, occur frequently, as it was expected. Whereas the wavelength limits and frequency limits of the modes depend on the distance between the surface and the D-layer, by the application of Spw-s it is possible to obtain a real and continuous monitoring of the fluctuation of the height of the lower border of the ionosphere for the whole Earth's surface, depending on the day, season, space-weather, seismic activity, and other global variations. The reasons of different changes are expected to be determined and classified using time series, and other significances resulted by continuous monitoring. This means one of the important research goals in the future.

Figure 9.
The generation mechanism and propagation path of spiky whistlers (SpW) [5]



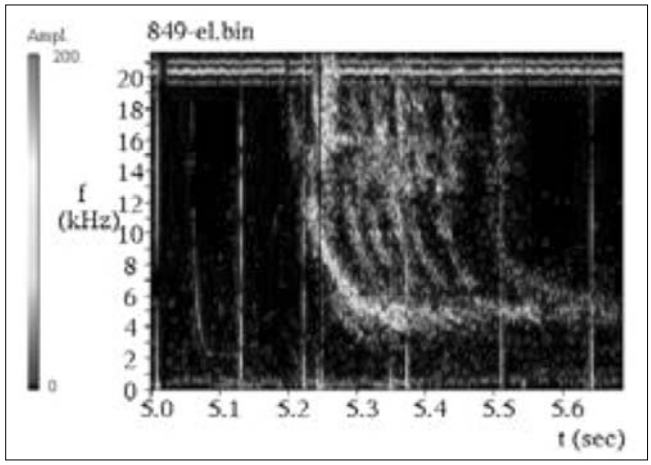
3.3 Direct verification for existence of guided whistlers

As it was mentioned in the Introduction, the appearance of whistlers at the Earth's surface opens some questions, as these signals are excited by lightning discharges and through the magnetosphere they reach the conjugate (northern or southern) hemisphere. The modeling of the connecting out/in process is important in the study and observation of the high atmosphere, in determination of the plasma parameters of the magnetized plasma surrounding the Earth, etc. Because of the fundamental importance of this question we have started to install the globally extended AWDA network, with terrestrial VR-1 and VR-2 measuring stations compatible with our onboard instruments, which can detect, classify and fully automatically evaluate the whistlers in 24 hours a day [9-11].

The early investigations based on calculations of strictly monochromatic (sinusoidal), continuous signals in the VLF band by ray tracing methods and these results suggested the hypothesis, that the whistlers can only and exclusively reach the conjugate point of the geomagnetic field line, if there is a conducting tubular structure (duct) along the magnetic line, formed by the inhomogeneous plasma density [7].

This is not impossible, as the charged particles in the Earth's atmosphere can easily move along the geomagnetic field lines, but this is hard for them into the

Figure 11.
SpW group detected by SAS2 on board of Compass-2 (16th of March, 2007)



rectangular direction. So an inhomogeneous density structure extends along the geomagnetic field lines in the magnetosphere.

But the large number of the whistlers would need the continuous presence of numerous “spaghetti”-like ducts with high stability. Another problem appears in the shape of the detected whistlers, which can be well described by free space propagation in anisotropic, magnetized plasma even in the approximate models or in the exact UWB models [1]. As a connection between the two possibilities, a shield-like structure of the ducting plasma inhomogeneities (“onion-skin”) can be also assumed in the high atmosphere, and this more stable and simpler structure ducts the whistlers along the magnetic field lines.

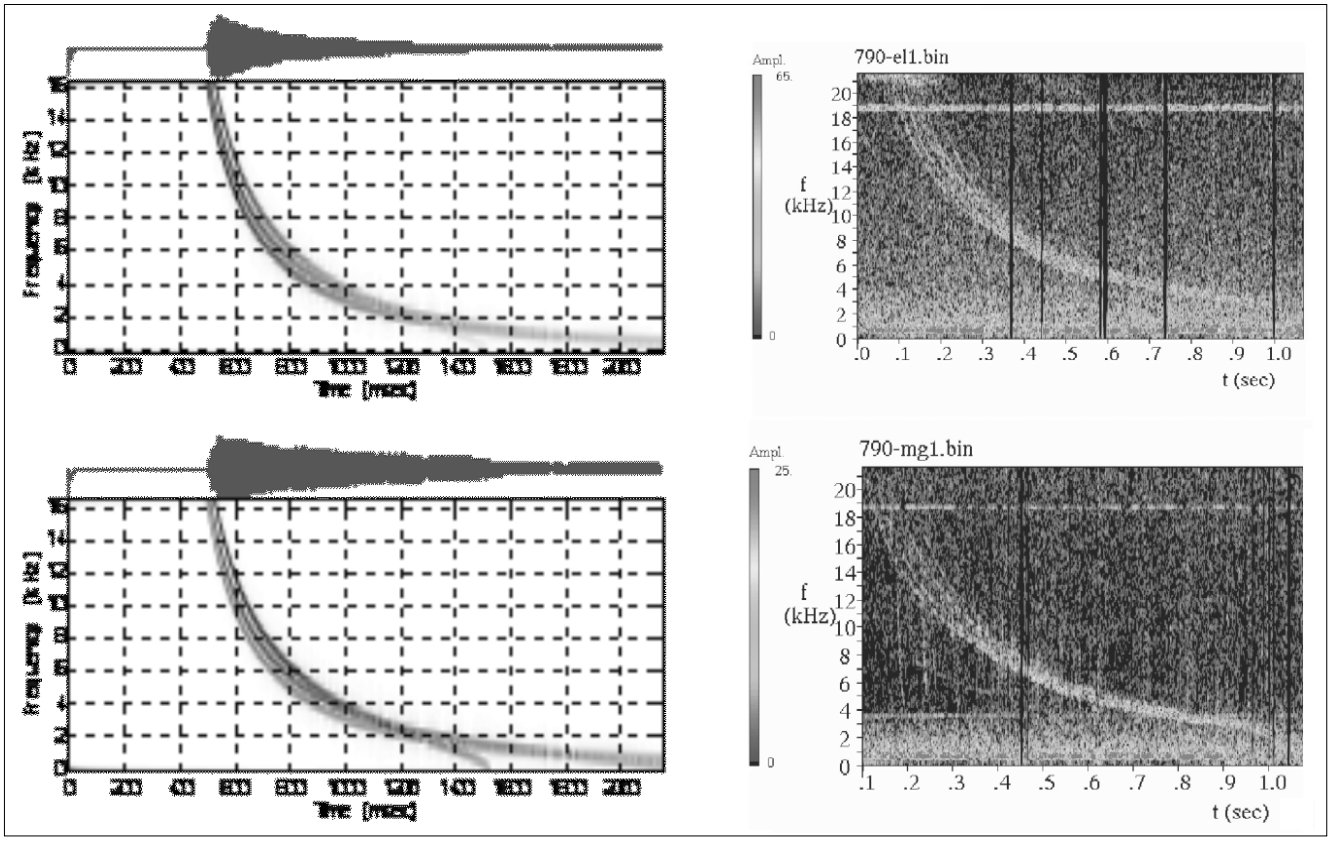
We successfully solved the propagation of UWB signals, impulses in rectangular waveguides filled by homogeneous, anisotropic, magnetized, cold plasma [4]. The results enlightened, on the one hand, that the wave pattern appearing during this propagation significantly differs from the free space propagation (as this was expected), on the other hand the shape of their spectrogram has no asymptotic limited wavelength, but the curvature of their spectral function turns to zero frequency in a finite time (this was unexpected). (Because of this unordinary result we have checked the model for two years, but no error was found in that.) This problem can only controlled by measurements, if the instrument can measure not only in the VLF, but in the ULF-ELF band also with high sensibility, and the convergence of the dynamic spectrum of the signal to zero can be well analyzed on the registered data. The SAS2-K2 fits well to these requirements.

A whistler group detected by Compass-2 on 28th of February, 2007 shows the form expected from our theoretical considerations (see *Fig. 12*). This is the first time, when a really ducted whistler can be presented. Using our theoretical model, we investigated the structure in which this signal form could appear. We can say on the basis of the calculations that the registered whistler propagated in a higher ducted mode along the magnetic field line. The ducting structure was not tubular, the signal propagated between surfaces, so the ducting structure was “onion-skin” like.

When we applied two surfaces in a 6 km distance from each other in our model calculation, the calculated and the measured data coincided with high accuracy. (The magnetic field is parallel with the bordering surfaces, and the direction of the propagation is parallel with the magnetic field.) The length of the propagation path is 30.000 km, the plasma frequency in the plasma model is 2.5 Mrad/s and the gyro-frequency is 900 krad/s.

The dynamic spectra of the first, second and third order modes can be seen in Fig. 12. As it can be seen in their dynamic spectra, the third order mode of the calculated signal coincides well with the measured signal. As a consequence of this result, it can be declared, that this signal is a ducted signal propagated in third mode

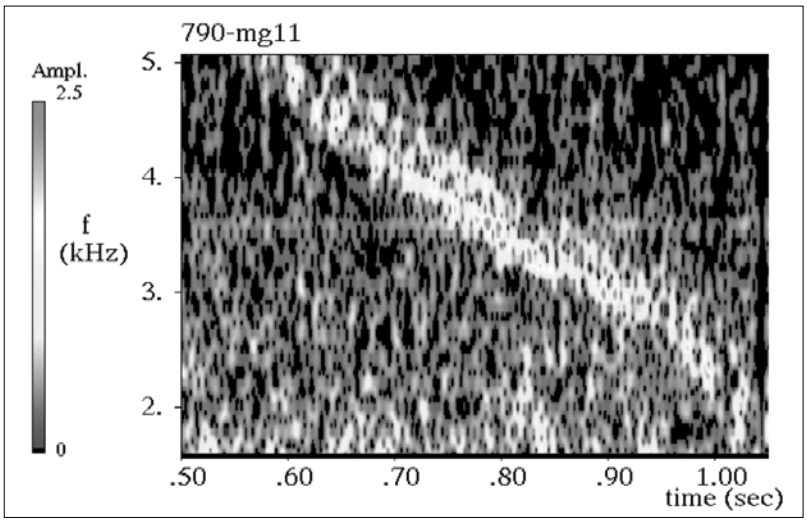
Figure 12.
First, second and third order modes of calculated UWB signals, in comparison with the dynamic spectrum of a whistler detected by Compass-2 on 28th of February, 2007



between two ducting surfaces. However, from these registrations the existence of the so-called ducts has not been verified, moreover, the presence of ducts is not indispensable for the propagation of whistlers. Furthermore it can be seen, that the whistlers having no significances of ducted propagation were not propagating in ducting structures. Because of this it is undoubtedly necessary to start intensive research in order to clarify the amount and the location of the ducted and the non-ducted propagation. (Reasonably, it is necessary to develop the UWB modeling of different propagation situations, too. This needs numerous theoretical breakthroughs as well.)

This unordinary, unusual signal form makes extensive, multilateral analysis and controlling necessary. Because of this we have examined the signal by digital matched filtering [6]. *Fig. 13* shows the FFT patters of the final section of the same signal, the matched filtered pattern of which can be seen in *Fig. 14*. The signal from and the method of its convergence to zero frequency (bending down) are unquestionable facts.

Figure 13.
FFT pattern of the final section of a measured ducted whistler



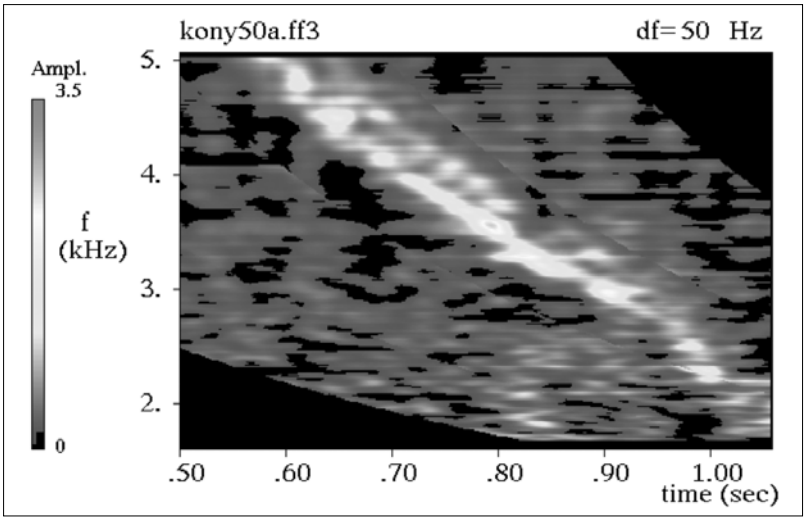
4. Summary

a) The technical probe of the Compass-2 and the onboard SAS2-K2 measuring system were successfully conducted.
b) The SAS2-K2 made it possible to get fundamentally new and important knowledge.
c) It is necessary to restart the investigation of the whistler-doublets, and to clarify the reason of the duplication in their spectra.
d) The SpW-s are applicable and effective tools in the continuous and global investigation of the dynamics of the lower border of the ionosphere.
e) We succeeded first time to demonstrate a ducted whistler signal, and to open the way for a more detailed investigation of the problem of ducted/non-ducted propagation, beyond the former theories on this topic.

Acknowledgement

The authors are greatly indebted to the Hungarian Space Office (Ministry of Environment and Water) in Hungary, to the IZMIRAN in Russia, and to the Lviv Centre of Institute of Space Sciences in Ukraine for the financial and other support of this R&D project.

Figure 14.
Matched filtered pattern of the final section of a ducted whistler



References

[1] Ferencz, Cs., Ferencz, O.E., Hamar, D. and Lichtenberger, J. (2001), Whistler phenomena. Short impulse propagation, Astrophysics and Space Science Library, Kluwer Academic Publishers, Dordrecht, NL.

[2] Ferencz, Cs., Ferencz, O.E., Lichtenberger, J., Székely, B., Steinbach, P. and Bodnár, L. (2006a), Élet egy csillag szomszédságában, (In Hungarian. "Life in the vicinity of a star") Híradástechnika, Vol. LXI., 2006/4, pp.29–33.

[3] Ferencz, O.E., Steinbach, P., Ferencz, Cs., Lichtenberger, J., Hamar, D., Berthelier, J.-J., Lefeuvre, F. and Parrot, M. (2006b), Full-wave modeling of long subionospheric propagation and fractional hop whistlers on electric field data of the DEMETER satellite, Int. Symp. Demeter, Toulouse, June 14-16, 2006.

[4] Ferencz, O.E., Steinbach, P., Ferencz, Cs., Lichtenberger, J., Parrot, M. and Lefeuvre, F. (2006c), UWB modeling of guided waves in anisotropic plasmas, 2nd VERSIM Workshop 2006, Sodankylä, Abtracts p.27, ISBN 951-42-6053-8.

[5] Ferencz, O.E., Ferencz, Cs., Steinbach, P., Lichtenberger, J., Hamar, D., Parrot, M., Lefeuvre, F. and Berthelier, J.-J. (2007), The effect of subionospheric propagation on whistlers recorded by the DEMETER satellite- observation and modeling, Annales Geophysicae, 25, pp.1103–1112.

[6] Hamar, D. and Tarcsai, Gy. (1982), High resolution frequency-time analysis of whistlers using digital matched filtering, Part I: Theory and simulation studies, Annales Geophysicae, 38, pp.119–128.

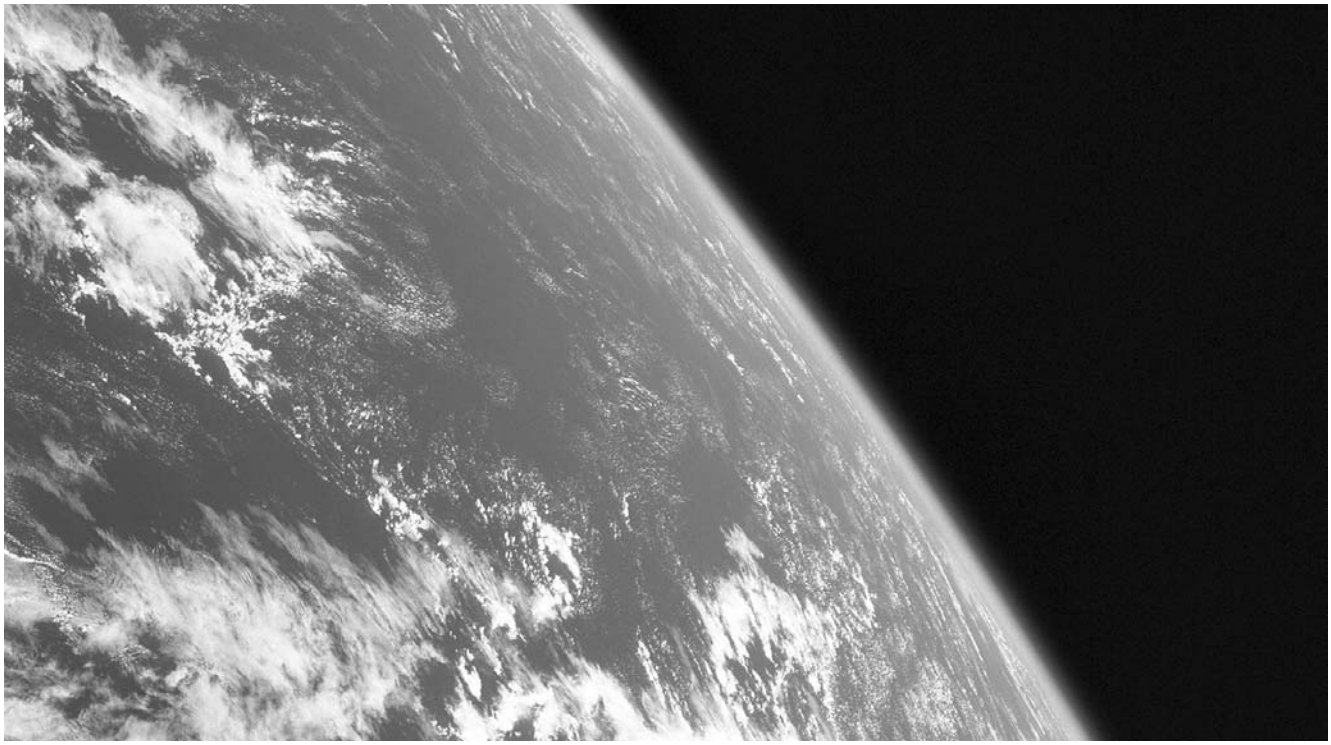
[7] Helliwell, R.A. (1965), Whistlers and related ionospheric phenomena, Stanford University Press, Stanford, California.

[8] Lichtenberger, J., Tarcsai, Gy., Pásztor, Sz., Ferencz, Cs., Hamar, D., Molchanov, O.A. and Golyavin, A.M. (1991), Whistler doublets and hyperfine structure recorded digitally by the Signal Analyzer and Sampler on the Active satellite, Journal of Geophysical Research, 96, 21,149-21, p.158.

[9] Lichtenberger, J., Bodnár L., Ferencz Cs., Ferencz O.E., Hamar D. and Steinbach P. (2001), Automatic whistler detector, First results, IAGA IASPEI, G2.07, Hanoi, Vietnam, p.124.

[10] Lichtenberger, J., Ferencz E.O., Bodnár, L., Ferencz, Cs. and Steinbach, P. (2006), Változóban a Föld-képünk, (In Hung. "Our picture from the Earth is changing") Híradástechnika, Vol. LXI., 2006/4, pp.51–53.

[11] Lichtenberger, J., Ferencz, Cs., Hamar, D., Steinbach, P. and Bodnár, L. (2007), Automatic whistler detector and analyzer system, Geophysical Research Abstracts, 6, 01390.



Hidden-Markov-Model based speech synthesis in Hungarian

BÁLINT TÓTH, GÉZA NÉMETH

Department of Telecommunications and Media Informatics, Budapest University of Technology and Economics

{toth.b, nemeth}@tmit.bme.hu

Keywords: speech synthesis, Text-To-Speech (TTS), Hidden-Markov-Modell (HMM)

This paper describes the first application of the Hidden-Markov-Model based text-to-speech synthesis method to the Hungarian language. The HMM synthesis has numerous favorable features: it can produce high quality speech from a small database, theoretically it allows us to change characteristics and style of the speaker and emotion expression may be trained with the system as well.

1. Introduction

There are several speech synthesis methods: articulatory and formant synthesis (trying to model the speech production mechanism), diphone and triphone based concatenative synthesis and corpus-based unit selection synthesis (based on recordings from a single speaker). Currently the best quality is produced by the unit selection method, although the size of the corpus database is quite big (it may be in the gigabyte range), voice characteristics are defined by the speech corpus and these features cannot be changed without additional recordings, labelling and processing which increase the cost of generating several voices.

Text-to-speech (TTS) systems based on the Hidden-Markov-Model (HMM) are categorized as a kind of unit selection speech synthesis, although in this case the units are not waveform samples, but spectral and prosody parameters extracted from the waveform. HMMs are responsible for selecting those parameters which most precisely represent the text to be read. A vocoder generates the synthesized voice from these parameters. HMM-based text-to-speech systems are becoming quite popular nowadays because of their advantages: they are able to produce intelligible, naturally sounding voice in good quality and the size of the database is small (1,5-2 Mbytes). It is also possible to adapt the voice characteristics to different speakers with short recordings (5-8 minutes) [1-4], and emotions can also be expressed with HMM TTS [5].

The current paper gives an overview about the architecture of HMM based speech synthesis, investigates the first adaptation of an open-source HMM-based TTS to Hungarian, and describes the steps of the adaptation process. The results of a MOS-like test are also introduced and future plans of the authors are mentioned as well.

2. The basics of Hidden-Markov-Models

Hidden-Markov-Models are mainly used for speech recognition [6] in speech research, although in the last decade or so they have also been applied to speech synthesis. The current section briefly introduces the basics of HMMs, a detailed description can be found in [7].

A HMM $\lambda(A,B,\pi)$ is defined by its parameters: A is the state transition probability, B is the output probability and π is the initial state probability. In case of text-to-speech let us assume that λ is a group of HMMs, which represent quintphones (a quintphone is five phones in a sequence) in sequence (*Fig. 1*). The series of quintphones define the word, which we would like to generate. The goal is to find the most probable state sequence of state feature vectors **X**, which will be used to generate the synthetic speech (see Section 3 for more details).

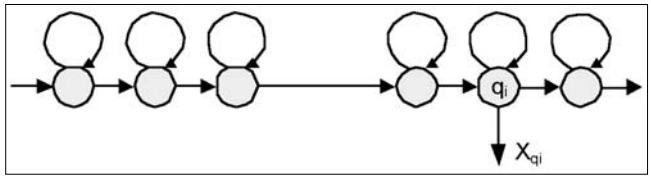


Figure 1. Concatenated HMM chain in q_i state, at i^{th} time, the output is X_{qi}.

The X_{qi} output is an M-dimensional feature vector at state q_i of model λ:

$$X_{q_i} = (x_1^{(q_i)}, x_2^{(q_i)}, x_3^{(q_i)} \ldots x_M^{(q_i)})^T$$

The aim is to define the $\underline{x} = (X_{q_1}, X_{q_2} \ldots X_{q_L})$ output feature vector (of length is L, which is the number of sounds/phonemes in an utterence) from model λ, which maximizes the overall likelihood $P(x|\lambda)$:

$$\underline{x} = \arg\max_x \{P(x|\lambda)\} = \arg\max_x \left\{ \sum_Q P(x|q,\lambda) P(q|\lambda) \right\},$$

Where $Q = (q_1, q_2 \ldots q_L)$ is the state sequence of model λ. The $P(x|\lambda)$ overall likelihood can be computed by adding the product of joint output probability $P(x|q,\lambda)$ and state sequence probability $P(q|\lambda)$ over all possible Q state-sequence.

To be able to compute the result within moderate time, the Viterbi-approximation is used:

$$\underline{x} \approx \arg\max_x \{P(x|q,\lambda,L) P(q|\lambda,L)\}$$

The $\underline{q}$ state sequence of model λ can be maximized independently of $\underline{x}$:

$$\underline{q} = \arg\max_{q} \{P(q|\lambda, L)\}$$

Let us assume that the output probability distribution of each state q_i is a Gaussian density function with μ_i mean value and Σ_i covariance matrix. The model λ is the set of all mean values and covariance matrices for all N states:

$$\lambda = (\mu_1, \Sigma_1, \mu_2, \Sigma_2 \ldots \mu_N, \Sigma_N)$$

Consequently the log-likelihood function is:

$$\ln\{P(x|q,\lambda)\} = -\frac{LM}{2}\ln\{2\pi\} - \frac{1}{2}\sum_{t=1}^{L}\ln\{|\Sigma_{q_t}|\} - \frac{1}{2}\sum_{t=1}^{L}(x_t - \mu_{q_t})^T \Sigma_{q_t}^{-1}(x_t - \mu_{q_t})$$

If we maximize x then the solution is trivial: the output feature vector equals to the states' mean values

$$\underline{x} = (\mu_{q_1}, \mu_{q_2} \ldots \mu_{q_L})$$

This solution is not representing the speech well because of discontinuities at the state boundaries. The feature vectors must be extended by the delta and delta-delta coefficients (first and second derivatives):

$$\underline{x} = ((x_{q_i})^T, (\Delta x_{q_i})^T, (\Delta^2 x_{q_i})^T)$$

3. HMM-based speech synthesis

HMM-based speech synthesis consists of two main phases: the training phase *(Fig. 2)* and the speech generation phase *(Fig. 3)*. During training the HMMs "learn" the spectral and prosodic features of the speech corpus, during speech generation the most likely parameters of the text to be synthesized are extracted from the HMMs.

For training a rather large speech corpus, the phonetic transcription of it and the precise position of the phoneme boundaries are required. The mel cepstrum, its first and second derivatives, the pitch, its first and second derivatives are extracted from the waveform. Then the phonetic transcription should be extended by context dependent labelling (see Subsection 4.2). When all these data are prepared, training can be started. During the training phase the HMMs are "learning" the spectral and excitation parameters according to the context dependent labels. To be able to model parameters with varying dimensions (e.g. $\log\{F_0\}$ in case of unvoiced regions) multidimensional probability density functions are used. Each HMM has a state duration density function to model the rhythm of the speech.

There are two standard ways to train the HMMs: (1) with a 2-4 hour long speech corpus from one speaker or (2) with speech corpora from more speakers and then adapt it to a speaker's voice characteristics with a 5-8 minute long speech corpus [1,2]. This way new voice characteristics can be easily prepared with a rather small speech corpus. According to [1,2] the adaptive training technique produces better quality than training from one speech corpus only. Furthermore there are numerous methods to change the voice characteristics [3,4].

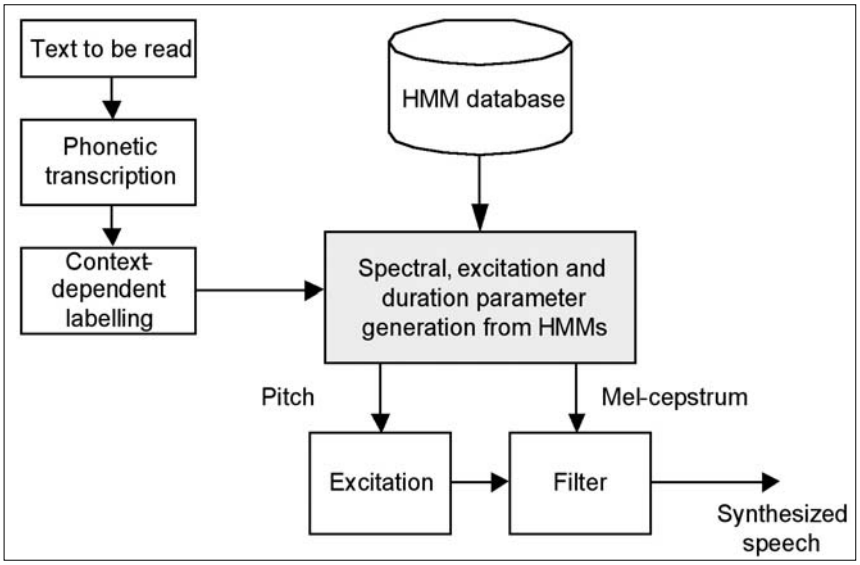


Figure 3. Speech generation with HMMs

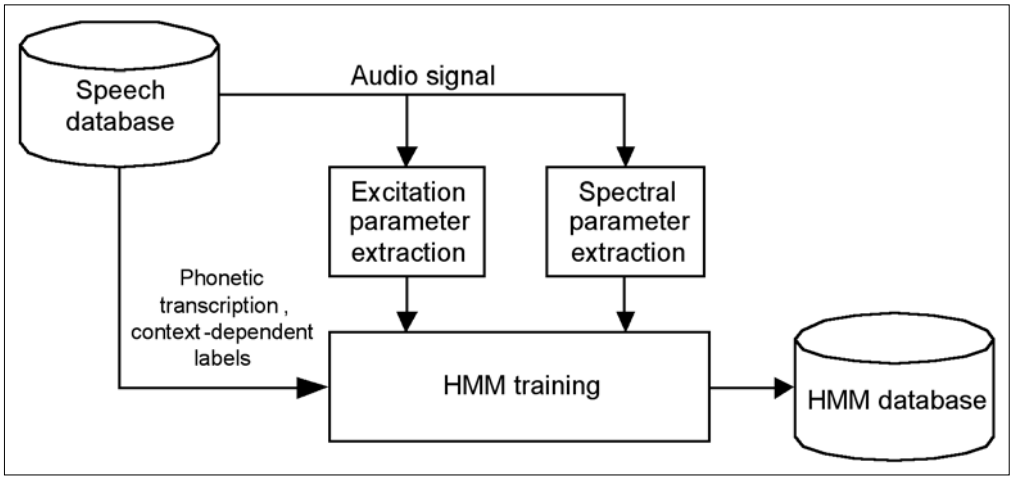


Figure 2. Training of the HMMs

Sounds	• The two previous and the two following sounds/phonemes (quintphones). Pauses are also marked.
Syllables	• Mark if the actual/previous/next syllable is accented. • The number of phonemes in the current/previous/next syllable. • The number of syllables from/to the previous/next accented syllable. • The vowel of the current syllable.
Word	• The number of syllables in the current/previous/next word. • The position of the current word in the current phrase (forward and backward).
Phrase	• The number of syllables in the current/previous/next phrase. • The position of the current phrase in the sentence (forward and backward).
Sentence	• The number of syllables in the current sentence. • The number of words in the current sentence. • The number of phrases in the current sentence.

Table 1.
The prosodic features used for context dependent labelling

To generate speech after training is done, the phonetic transcription and the context dependent labelling (see Subsection 4.2) of the text to be read must be created. Then the phones' duration is extracted from the state duration density functions and the most likely spectral and excitation parameters are calculated by the HMMs. With these parameters the synthetic speech can be generated: from the pitch value the excitation signal of voiced sections is generated and then it is filtered, typically with a mel log spectrum approximation (MLSA) filter [8]. Simple vocoders were used earlier, lately mixed excitation models are applied in order to achieve better quality [9].

4. Adapting HMM-based TTS to Hungarian

The authors conducted the experiments with the HTS framework [10]. For the Hungarian adaptation a speech database, the phonetic transcription of it, the context-dependent labelling and language specific questions for the decision trees were necessary. In the following the most important steps will be described.

4.1 Preparation of the speech corpus

The authors used 600 sentences to train the HMMs. All the sentences were recorded from professional speakers, it was resampled at 16.000 Hz with 16 bits resolution. The content of the sentences is weather forecast and the total length is about 2 hours. The authors prepared the phonetic transcription of the sentences, the letter and word boundaries were determined by automatic methods, which are described in [11].

4.2 Context-dependent labelling

To be able to select to most likely units, a number of phonetic features should be defined. These features are calculated for every sound.

Table 1 summarizes the most important features.

Labelling is done automatically, which may include small errors (e.g. defining the accented syllables), although it does not influence the quality much, as the same algorithm is used during speech generation, thus the parameters will be chosen by the HMMs consistently (even in case of small errors in the labelling).

4.3 Decision trees

In Subsection 4.2 context-dependent labelling was introduced. The combination of all the context-dependent features is a very large number. If we take into account the possible variations of quintphones only, even that is over 160 million and this number increases exponentially if other context dependent features are included as well. Consequently, it is impossible to design a natural speech corpus, where all the combinations of context-dependent features are included. To overcome this problem, decision tree based clustering [12,13] is used.

As different features influence the spectral parameters, the pitch values and the state durations, decision trees must be handled separately. *Table 2* shows which features were used in case of Hungarian for the decision trees [14].

For example, if the decision tree question regarding the length of the consonants is excluded, then the HMMs will mostly select short consonants even for long ones, as these are not clustered separately and there are much more short consonants in the database.

4.4 Results

In order to be able to define the quality of Hungarian HMM-based text-to-speech synthesis objectively, the authors conducted a MOS (Mean Opinion Score) like listening test. Three synthesis engines were included in the test: a triphone-based, a unit selection system and a HMM-based speech synthesis engine.

Table 2.
Features used for building the decision trees

Phonemes	• vowel / consonant • short / long • stop / fricative / affricative / liquid / nasal • front / central / back vowel • high / medium / low vowel • rounded / unrounded vowel	
Syllable	• stressed / not stressed • numeric parameters	*(see Table 1.)*
Word	• numeric parameters	*(see Table 1.)*
Phrase	• numeric parameters	*(see Table 1.)*
Sentence	• numeric parameters	*(see Table 1.)*

At the beginning of the test 3-3 sentences randomly generated from each system were played, these sentences were not scored by the subjects. The reason for doing this is to have the subjects used to synthetic voice and to show them what kind of qualities they can expect.

At the next step 29 sentences generated by each system were randomly played in different sequences in order to avoid 'memory effects' [15]. The content of the test sentences were from the weather forecast domain. The triphone based system is the ProfiVox domain independent speech synthesizer, the HMM based TTS was trained with weather forecast sentences (cca. 600 sentences) and the unit selection system had a large weather forecast speech corpus (including cca. 7000 sentences). The same 29 sentences were generated by all systems, but none of these sentences were present in the speech corpora. The subjects scored the sentences from 1 to 5 (1 was the worst, 5 was the best).

12 subjects were included in the test. The results are shown in *Fig. 4*.

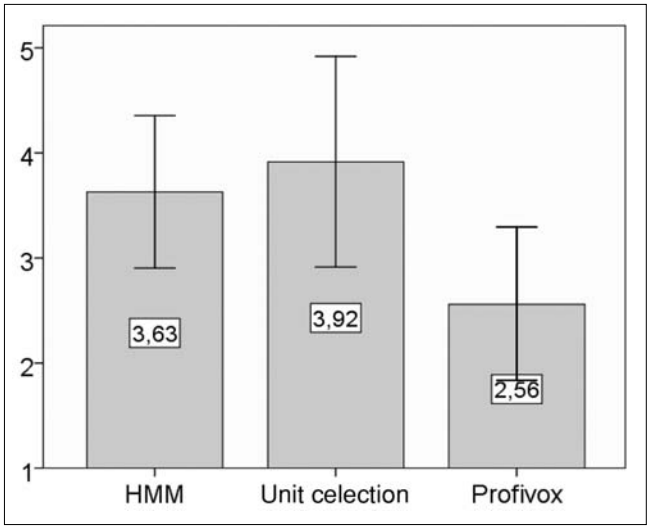


Figure 4.
The results of the MOS like test

The triphone-based system scored 2.56, the HMM-based TTS scored 3.63 and the unit selection one scored 3.9 on average (mean value). The standard deviation was in the same order 0.73, 1 and 0.73. Although the unit selection system was considered better then the HMM-based TTS, it can read only domain-specific sentences with the same quality, while the HMM-based TTS can read any sentence with rather good quality. Furthermore the database of the unit selection system includes more then 11 hours of recordings, while only 1.5 hours of recordings was enough to train the HMMs. The size of the HMM database is under 2 Mbytes, while the unit selection system's database is over 1 GByte.

The triphone based system was designed for any text, domain specific information is not included in the engine. This can be one reason for the lower score. The absolute value of the results is not so important, rather the ratio of them contains the most relevant information.

5. Future plans

The current paper introduced the first version of the Hungarian HMM-based text-to-speech system. As the next step the authors would like to record additional speech corpora in order to test adaptive training, to achieve more natural voice and to be able to create new voice characteristics and emotional speech with small (5-8 minutes long) databases.

Because of the small footprint and the good voice quality, the authors would like apply the system on mobile devices as well. To be able to port the hts_engine to embedded devices, it may occur, that the core engine must be optimized to achieve real-time response on mobile devices.

6. Summary

In this paper the basics of Hidden-Markov-Models were introduced, the main steps of creating the first Hungarian HMM-based text-to-speech synthesis system were described and a MOS-like test was presented. The main advantage of HMM-based TTS systems is that they can produce natural sounding voice from a small database, it is possible to change the voice characteristics and to express emotions.

Acknowledgements

We would like to thank the test subjects for their participation. We acknowledge the support of Mátyás Bartalis for creating the Web-based MOS-like test environment and the help of Péter Mihajlik in using the Hungarian speech recognition tools. The research was partly supported by the NKTH in the framework of the NAP project (OMFB-00736/2005).

Authors

BÁLINT PÁL TÓTH graduated in 2005 and received a diploma with honours from the Budapest University of Technology and Economics, Department of Telecommunications and Media Informatics. He continued his research on speech synthesis and multimodal user interfaces as Ph.D. student right after the diploma. His main research topics are Hidden Markov-Model based speech synthesis and multimodal user interfaces on mobile devices.

GÉZA NÉMETH graduated from the Budapest University of Technology and Economics, Faculty of Electrical Engineering, in 1983, and obtained an engineering specialization diploma in 1985. He worked as a development engineer in BEAG Electroacoustic Factory between 1985 and 1987. Dr. Németh has been with the Department of Telecommunications and Media Informatics at Budapest University of Technology and Economics since 1987 where has been teaching measurement technologies, communication systems, telecommunications, signal processing, telecommunication management, speech information systems. He is also leading the Speech Technology Laboratory. Dr. Németh has been instrumental in transferring speech research results into practice, several applications have been developed under his leadership.

References

[1] T. Masuko, K. Tokuda, T. Kobayashi, S. Imai, "Voice characteristics conversion for HMM-based speech synthesis system," Proceedings of ICASSP, 1997, pp.1611–1614.

[2] M. Tamura, T. Masuko, K. Tokuda, T. Kobayashi, "Adaptation of pitch and spectrum for HMM-based speech synthesis using MLLR," Proceedings of ICASSP, 2001, pp.805–808.

[3] T. Yoshimura, K. Tokuda, T. Masuko, T. Kobayashi, T. Kitamura, "Speaker interpolation in HMM-based speech synthesis system," Proceedings of Eurospeech, 1997, pp.2523–2526.

[4] M. Tachibana, J. Yamagishi, T. Masuko, T. Kobayashi, "Speech synthesis with various emotional expressions and speaking styles by style interpolation and morphing," IEICE Trans. Inf. & Syst., Vol. E88-D, No.11, 2005. pp.2484–2491.

[5] S. Krstulovic, A. Hunecke, M. Schroeder, "An HMM-Based Speech Synthesis System applied to German and its Adaptation to a Limited Set of Expressive Football Announcements," Proceedings of Interspeech, 2007.

[6] Mihajlik P., Fegyó T., Németh B., Tüske Z., Trón V., "Towards Automatic Transcription of Large Spoken Archives in Agglutinating Languages: Hungarian ASR for the MALACH Project," In: Matousek V., Mautner P. (eds.), Text, Speech and Dialogue: 10th International Conf., TSD 2007, Pilsen, Czech Republic, September 2007. Proceedings, Berlin; Heidelberg: Springer, Lectures Notes in Computer Sciences, pp.342–350. (Lecture Notes in Artificial Intelligence, p.4629.)

[7] Lawrence R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," Proceedings of the IEEE, 77 (2), Febr. 1989, pp.257–286.

[8] S. Imai, "Cepstral analysis synthesis on the mel frequency scale" Proceedings of ICASSP, 1983, pp.93–96.

[9] R. Maia, T. Toda, H. Zen, Y. Nankaku, K. Tokuda, "A trainable excitation model for HMM-based speech synthesis," Proceedings of Interspeech, Aug. 2007, pp.1909–1912.

[10] H. Zen, T. Nose, J. Yamagishi, S. Sako, T. Masuko, A.W. Black, K. Tokuda, "The HMM-based speech synthesis system v.2.0," Proceedings of ISCA SSW6, Bonn, Germany, Aug. 2007.

[11] Mihajlik, P., Révész, T., Tatai, P., "Phonetic transcription in automatic speech recognition," Acta Linguistica Hungarica, 2003, Vol. 49., No.3/4, pp.407–425.

[12] J. J. Odell, "The Use of Context in Large Vocabulary Speech Recognition," PhD dissertation, Cambridge University, 1995.

[13] K. Shinoda and T. Watanabe, "MDL-based context-dependent subword modeling for speech recognition," J. Acoust. Soc. Jpn. (E), Vol. 21., No.2, 2000. pp.79–86.

[14] Gósy M., Fonetika, a beszéd tudománya, Budapest, Osiris Kiadó, 2004.

[15] Jan P. H. van Santen, Perceptual experiments for diagnostic testing of text-to-speech systems, Computer Speech and Language, 1993, pp.49–100.

Using prosody for the improvement of automatic speech recognition

GYÖRGY SZASZÁK, KLÁRA VICSI

Department for Telecommunication and Media Informatics,
Budapest University for Technology and Economics

{szaszak, vicsi}@tmit.bme.hu

Keywords: speech recognition, prosody, boundary detection, prosodic segmentation

This paper describes sentence, phrase and word boundary detection based on prosodic features, implemented in a HMM-based prosodic segmentation tool. Integrated into a speech recognizer, an N-best rescoring is performed based on the output of the prosodic segmenter, which determines the prosodic structure of the utterance. In an ultrasonography task, we obtained 3,82% speech recognition error reduction using a simplified bi-gram language model.

1. Introduction

Prosody or supra-segmental features are integrant parts of human speech, they provide cues for the listener to understand the meaning by segmenting the speech flow, by emphasizing the important or new information, etc. Moreover, prosody carries sentence mood (modality) and allows the speaker to express emotions, which are embedded acoustically into the speech utterance.

From the point of view of speech technology, high quality speech synthesis would be impossible without modelling prosody, which means definition of the proper intonation, stress and logical segmentation. In speech recognition, however, prosody was not addressed as an information source for a long time, even if supra-segmental features provide not only segmentation information or some representation of nuances in the meaning, but they might by themselves carry information not contained in any other speech related feature. Automatic speech recognizers should exploit this information source in order to ensure some redundancy for speech decoding and also to catch information which would be lost otherwise. For example, automatic classification of sentence modality can be crucial in several speech technology based information retrieval systems, hence several sentences can be composed from identical word chains, the meaning being still different because of the differing sentence mood [1] (question or statement, for example). This is even more important if – like in Hungarian – the subject-predicate inversion does not appear to predict syntactically the sentence modality. In traditional statistical speech recognition, sentence modality classification would be impossible in many cases.

Prosody can, however, be very useful also in traditional speech recognition by providing segmentation information (boundary detection) about the speech utterance. Boundaries of sentences, clauses, syntagms or even some word boundaries can be identified based on supra-segmental features, and the information about the temporal localization of these boundaries can help reduce searching space during the decoding process by removing or penalizing hypotheses not fitting the determined prosodic pattern. Searching space reduction means more robust (more accurate) and faster recognition, recognition speed being one of the critical factor when treating agglutinating languages like Hungarian, Finnish, Turkish, etc. in systems, if real time operation is a basic requirement.

Prosody can also help syntactical and semantic analysis [3] and can predict information-rich segments of speech by detecting stress.

Prosodic features – even if they have not became integral parts of speech recognizers yet – were examined and exploited by several research groups, mainly for English and German languages. Veilleux and Ostendorf elaborated an algorithm rescoring N-best lattices based on prosodic information [10]. N-best lattices are graphs representing recognition hypotheses, each arc having an associated score which functions as a weight, calculated from acoustic and linguistic analysis of the input speech. Based on prosodic information and analysis, these scores can be modified, this is called N-best rescoring. Indeed, it has the same effect as if a prosodic analyser module added his own scores to the acoustic and linguistic ones. The final recognition result is given as the path having the highest score (the most probable path) through the lattice. A similar work was presented for German language in [2].

Gallwitz et al. developed an integrated speech recognizer [1], treating and exploiting "traditional" acoustic and prosodic-acoustic features in parallel. The authors of the present article have also examined the use of prosody in speech recognition [12].

2. Extracting acoustic-prosodic information from speech

For representation of prosody, fundamental frequency (F0), energy level and time course are measurable. Based on our earlier analysis reported in [8], F0 and energy were found to be characteristic when considering em-

phasis detection. The extraction of prosodic information is performed using the Snack package of KTH [7]. The extraction of F0 is done by AMDF method using a 25 ms long window. The frame rate is set to 10 ms. The obtained F0 contour was firstly filtered with our anti-octave jump tool. This tool eliminates frequency halving and doubling, and also cuts F0 values associated to the first and last frames of each voiced speech segment. This was followed by a smoothing with a 5 point mean filter (5 points cover a window of about 50 ms) and then the log values of F0 were taken, which were linearly interpolated. During the interpolation, two restrictions must be fulfilled. Firstly, interpolation should not affect pauses in F0 longer than 250 ms; secondly, interpolation should be omitted if the initial value of F0 after an F0-gap higher than a threshold value. This threshold value depends on the last measured F0 values and equals the 110% of the average F0 value of the three last voiced frames before the gap (unvoiced period).

These restrictions affecting the interpolation were found necessary because an unvoiced period of length more than 250 ms is likely to be a silence, which should also be detected. On the other hand, interpolation of such a long period would yield only a broad approximation. The reason for the maximal rise criteria of 10% for F0 can be explained in the same manner: firstly a silence (including a breath) is likely, secondly, emphasis is expected to produce also such a rise which should not be smoothed by the interpolation. The threshold values to trigger interpolation were determined empirically. An automatic algorithm for the determination of these values based on speaker specific variables (such as speech or articulation rate, F0 dynamic range, etc.) would also be of interest in the future, but this problem is not issued in the current article.

Energy level values were also extracted using the Snack package, the window size (25 ms) and frame rate (10 ms) were identical to those applied for F0. Energy contour was then filtered by a mean filter. Unlike F0, energy level is a continuous variable, so interpolation is not necessary.

After feature extraction and basic shape conditioning described above, delta and acceleration coefficients are appended to both F0 and intensity streams. These coefficients are computed with a regression-based formula (1). The regression is performed in 3 different steps with increasing regression window length: firstly with a window of ±10 frames, secondly with a window of ±25 frames and finally, a window of ±50 frames is used (W in equation (1)). This means that the final feature vector consists of 14 elements (original F0 and intensity data + 3-3 delta + 3-3 acceleration components for both of them).

The formula applied was [9]:

$$d_t = \frac{\sum_{i=1}^{W} i(c_{t+i} - c_{t-i})}{2\sum_{i=1}^{W} i^2}, \quad (1)$$

where d_t is the delta coefficient at time t c_{t-i} and c_{t-i} are coefficients from the stream to be derivated, W is the window length given in the number of frames.

3. Using the prosodic information in the speech recognition process

Prosodic information is used to obtain a broad segmentation of the speech on sentence, clause, syntagm and word boundaries. Feeding this information into the speech recognizer, we except a higher accuracy and the implementation of functions presented in the introduction.

Our algorithm is based on the fact that stress in Hungarian is fixed [2]: if a word is stressed, stress is produced on the first syllable. This makes it possible to handle prosodic information without knowing the underlying word and phoneme sequence. Of course, the final aim is to integrate the processing of phoneme characteristic spectral and prosody affected syntactical information in the speech recognizer.

3.1 Training of an automatic prosodic segmenter

The prosodic segmentation is based on the intonation shape of individual stressed speech segments, separated on word boundaries. As a by-product of this recognition, the temporal location of these boundaries is also available. Boundaries are expected to occur on word boundaries, some of which can also be syntagm and/or clause and/or sentence boundaries at the same time. Please note that intonation now is defined in a more detailed interval than one sentence, as the intonation of the sentence is further split into intonationally coherent segments, so that they coincide with stress and hence by word boundaries. Further in the article, intonation is always regarded as some type of "sentence sub-intonation".

When determining the set of intonation types, a crucial step is to define classes which are well distinguishable and cover all frequent intonation patterns. To accomplish this, only 6 types of intonation patterns were defined. Silence is the 7th class. The used intonation pattern are listed in *Table 1*.

For training the prosodic segmenter, training samples were selected from Hungarian BABEL speech database [6] (22 speakers, 1600 sentences). This material was segmented based on intonation patterns shown in Table 1. An initial hand-labelling was then extended to

Table 1.
Intonation patterns used for prosodic segmentation

Label	Intonation pattern	Note
me	variable	Sentence onset unit
fe	rise (stress) – falling	Strongly stressed syntactical unit
fs	rise-fall	Stressed unit
mv	falling	Low sentence ending
fv	rising	High sentence or phrase ending
s	floating	Unstressed unit
sil	–	Silence

a computer aided segmentation using a primitive prosodic segmenter trained on hand-labelled data. Hand-labelling was performed relying on F0 and energy contour and subjective impression after listening.

The prosodic segmenter itself is a Markov-model based system whose structure is very close to standard HMM speech recognizers. The 14 dimensional acoustic-prosodic frames are calculated every 10 ms. The number of states is 11 (after optimization) for each intonation pattern class, the linear HMM models were implemented using the HTK package [9].

3.2 Prosodic segmentation process

Automatic prosodic segmentation is carried out using the same algorithms as in speech recognition: first, the acoustic pre-processing is performed, in our case, this is a prosodic-acoustic pre-processing as described in Section 2; then in the decoding stage, Viterbi algorithm is used to obtain the most probable intonation pattern sequence. Hence we use only 7 different pattern classes, and prosodic-acoustic observation vectors are only 14 dimensional and 1 or 2 Gaussians are sufficient for acoustic-prosodic modelling, the decoding process is very fast.

Similarly to a language model in speech recognition, a prosodic grammar is introduced for prosodic segmentation, which specifies the acceptable intonation pattern sequences. This prosodic grammar is relatively severe, but we found empirically that this improves significantly prosodic segmentation performance, while the number of cases where an error occurs due to insufficient generalization capabilities of the prosodic grammar is very low.

The prosodic grammar is given as (using notations from HTK Book [9], p.163):

$$Sentence = [sil] < [me]\ \{fe \mid fv\ [s]\ \}\ [mv]\ [sil] > sil \qquad (2)$$

Here, '<>' symbol pair refers to one or more, '{}' symbol pair to zero or more repetitions. The '|' symbol denotes alternatives, the '[]' pair encloses optional events. This proto-sequence is interpreted as the prosodic model of a sentence built from intonation patterns.

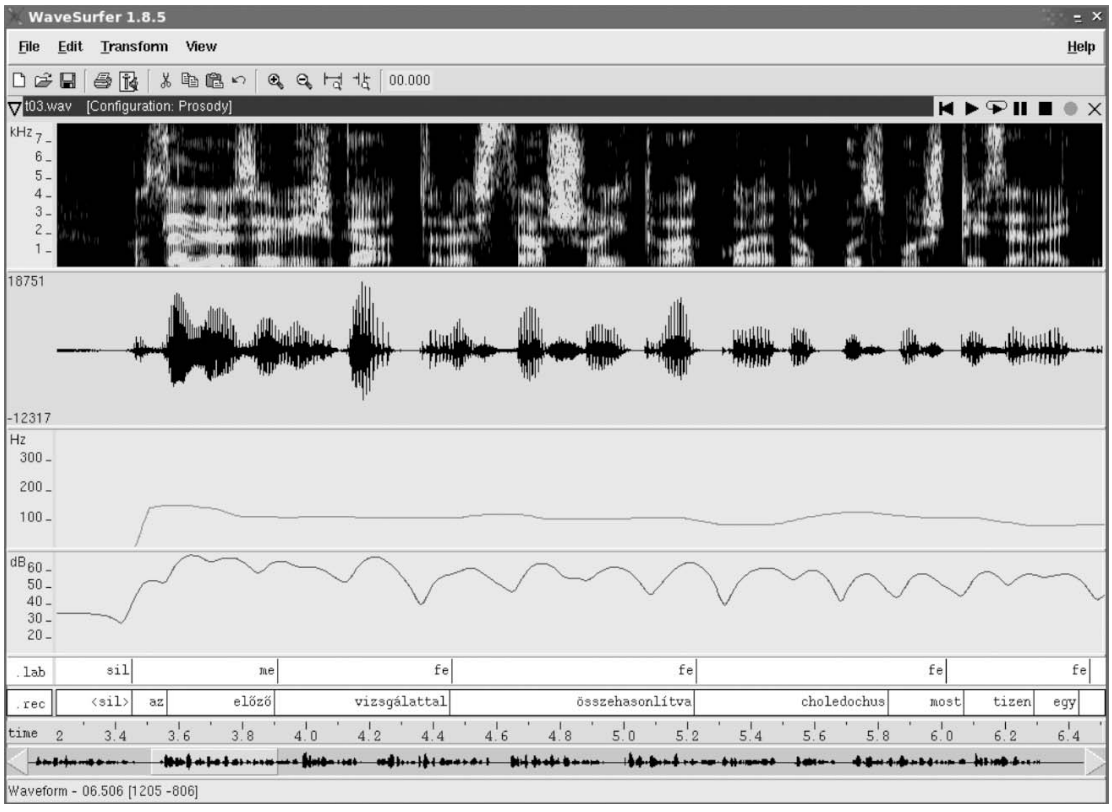
As a by-product of prosodic pattern alignment, the start and ending times of intonation pattern are also calculated. The example shown in *Fig. 1* illustrates the result of the prosodic segmentation process.

3.3 Integration of the prosodic segmenter into the speech recognizer

The output of the prosodic segmenter can be used in speech recognizers to obtain more accurate results and to reduce the searching space. Speech recognizers usually construct a graph (lattice) which specifies the possible outcomes (hypotheses) of the recognition process. Each arc in the graph has its own associated scores (weights) based on a calculation of acoustic and linguistic likelihoods given the input speech signal.

These scores can be re-evaluated (rescoring) with the prosodic information, and so the final recognition result (text output) takes into account prosodic characteristics of the speech. The rescored lattice then goes through the same parsing process as in a standard speech recognizer.

Figure 1. Result of prosodic segmentation for the Hungarian sentence "Az előző vizsgálattal összehasonlítva a choledochus most 11 milliméteres..."



Bounds in the figure from the top to down represent spectrogram (1), waveform (2), interpolated F0 (3), energy (4), prosodic segmentation (5) and underlying word sequence (6).

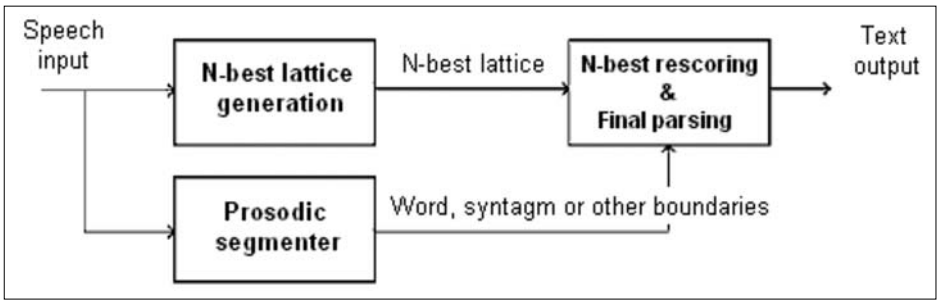


Figure 2.
Structure of a speech recognizer with prosodic module

The general recognition process with prosodic module is illustrated in *Fig. 2*.

3.4 Rescoring of N-best lattices

As briefly shown earlier, rescoring of N-best lattices is based on prosodic segmentation. The basic idea is that words or word chains (all recoverable from the N-best lattice) whose syntactical boundaries match well the prosodic structure defined by the prosodic segmentation should be promoted, this means the increase of their associated scores. Similarly, if the temporal characteristics found in the lattice do not fit the prosodic segmentation, the scores can be decreased.

However, the prosodic segmentation might also contain some errors. In spontaneous speech, several characteristic phenomena can lead to even higher prosodic segmentation error rates: mispronounciations, self-corrections, altered prosody or intensive emotions can all disturb the operation of automatic prosodic segmentation. We have already presented a detailed error analysis of the prosodic segmentation in our earlier work [8], now it is sufficient to remember that prosodically predicted boundaries should also be treated carefully when performing lattice rescoring.

As prosodic information is available in the supra-segmental domain, its time resolution is broader than that of the word or phoneme boundaries predicted by the speech recognizer itself. To illustrate this, let's have a look at a final unvoiced fricative of a word: our reference point in prosody is always the last voiced sound (vowel), this uncertainty about F0 is than around the length of a phoneme.

To overcome such difficulties, the locations of syntactic (sentence, phrase, syntagm or word) boundaries (t_B) are transformed to intervals to allow some ΔT time shift when aligning prosodic segmentation to the lattice. Within this interval, the boundary likelihood (L_B) is the highest in the middle and is decreasing towards the limits as defined by:

$$L_B(t) = \begin{cases} A\cos\left(\frac{\pi}{2\Delta T}t\right) + C, \; if \; t \in \left[t_B - \Delta T, t_B + \Delta T\right] \\ 0 \quad otherwise \end{cases} \tag{3}$$

where A and C are constants. (In our experiments to be presented in Section 4, ΔT was set to 10 frames, which equal 100 ms.) The cosine function was chosen for its simplicity, as it is required that the point to interval transform function has a flat maximum at t_B and decreases towards the limits of the ΔT interval.

The N-best lattice rescoring is then performed as follows. Each edge in the lattice has a word or a word chain associated (with a combined acoustic and linguistic score) and each node has its associated timestamp corresponding to the start and ending times of the word (chain) defined by the edges. A prosodic score is calculated based on the L_B curve, which is the higher if the actual node is the closer (see also Equation 3):

$$Sc_{renum} = w_a L_B(t_{start}) + w_b L_B(t_{end}), \tag{4}$$

where t_{start} is the timestamp of the start node of the word (chain) and t_{end} corresponds to the timestamp of the end node. w_a and w_b are weights.

Hereafter, $L_B(t_i)$ is summed for each frame i of the word (chain) – except the first and last k ones, where t_i is the time index of the actual frame:

$$Sc_{punish} = \sum_{i=k+1}^{N-k-1} L_B(t_i), \tag{5}$$

where N is the total number of frames associated to the word (chain), $k = \Delta T = 100$ ms.

The new $Sc_{rescored}$ score of the edge (and so of the word (chain)) is:

$$Sc_{rescored} = w_O Sc_{orig} + w_P (Sc_{renum} - Sc_{punish}), \tag{6}$$

where

Sc_{orig} is the original score, w_O and w_P are weights.

4. Experiment: integrating the prosodic segmenter into an ultrasonography speech recognizer

This section presents an experiment in which the prosodic segmenter functioned as part of a speech recognizer. The integration of the prosodic segmenter into the speech recognizer was carried out as presented in Section 3.3, the operation of the system was the same described in Section 3.4.

The speech recognizer was a Hungarian language, continuous speech recognizer with a 4000 word abdominal ultrasonography dictionary and a corresponding bi-gram language model. This latter was binarized, so it reflected only whether a word sequence was grammatically allowed or forbidden. This reduction was used in order to test the impact that prosodic information can add to speech recognition. However, in large vocabulary speech recognizers such a language model simplification can be useful, as the creation of a language model which covers representatively the application domain is

very time and money consuming, mainly for agglutinating languages – like Hungarian –, where even a relatively close application domain needs a larger vocabulary due to the several inflected forms of basic words.

The ultrasonography speech recognizer was implemented in HTK environment, using the "classical" 39 MFC coefficients, 32 Gaussian mixtures for each phoneme state and 10 ms frame rate. For training the 37 acoustic phoneme models, approx. 8 hours of speech was used form MRBA [11] database. The training corpus was segmented on phoneme level.

We integrated the prosodic segmenter into this recognizer in order to analyse recognition performance. The weights in equations (4) and (6) were set as follows: w_a=0,5, w_b=0,5, w_O=1, w_P=2,5.

4.1 Results

The testing was carried out on a set of 20 medical reports in the domain of abdominal ultrasonography. (A report contains approx 10 to 20 sentences.) The baseline and the integrated systems worked in an identical environment (same conditions, same recorded reports). Results are presented in *Table 2*. Out of 20, 6 medical reports were representatively selected to be presented in Table 2 in order to allow deeper analysis of results. The overall relative increase in the number of correctly recognized words was 3.82% for the whole test set.

The relative change in the number of correctly recognized words varies from report to report. In case of report ID 03, the relative improvement was over 10%, however, performance might be the same (ID 08) or even worse (ID 16) in the integrated prosodic segmenter system than in the baseline system. Further investigating each medical report and their prosodic segmentation, it was found that a decrease in the performance of the integrated system compared to the baseline one was caused by the errors of the prosodic segmenter, which can be misled by a less proper pronunciation in terms of supra-segmental features, or the error of the pitch detector algorithm can also lead to false boundary detection (prosodic segmentation).

Pitch detectors are sensible to hoarsed (glottalized) speech, some errors were also caused by this phenomenon. On the other hand, reports which were correctly uttered concerning prosody, show a higher improvement compared to the baseline system. A prosodically correct utterance does not require per se professional voicing skills, a common, prosodically well formed pronunciation is sufficient.

Table 2.
Ratio of correctly recognized words with baseline system vs. integrated system

Report ID	Correct words [%]		Relative change in # of correct words [%]
	Baseline	Integrated	
03	71,2	78,9	10,9
07	78,8	80,6	3,6
08	84,6	84,6	0,0
10	70,8	72,2	2,0
16	68,3	66,7	-2,4
19	83,8	90,5	8,1
Overall *(20 reports)*	***75,99***	***78,89***	***3,82***

Please note that in our algorithm, syntactical boundaries missed by the prosodic segmenter do not alter recognition performance. Of course, the more syntactic boundaries the prosodic segmentation reveals, the more performance improvement one can expect. Prosodic segmentation will never locate all of the word boundaries within the speech based solely on supra-segmental features, such a task would exceed even humans' capabilities.

This is why we used rather the *syntactical boundary* expression through the article instead of *word boundary,* but note also that a syntactical boundary is always a word boundary. We regard as proved that word boundary detection based on prosodic features can improve speech recognition performance.

As a general remark, we think that prosodic segmentation is not always as accurate in the temporal domain and in its resolution capabilities as it would be the ideal one to locate syntactic boundaries. However, this problem can be solved by tracking of the phoneme sequence which would allow a compensation of the prosodic structure in case of necessity. (Of course, tracking in speech recognition is always back-tracking with some delay.) For example, we have mentioned in Section 3.4 that unvoiced phonemes at the end of words can evoke an uncertainness concerning the F0 curve. Such a problem could be more efficiently treated if we knew the underlying phoneme structure or at least if we calculated some confidence of the prosodic segmentation based on phoneme context. We are planning to extend our research in this direction in the future.

5. Summary

Our article addressed the use of supra-segmental (or in other words prosodic) features in speech recognition. We have presented a prosodic segmenter, which aligns syntactical unit assigned intonation patterns or silence to the speech signal. Integrated into an automatic speech recognizer, the prosodic segmenter is used to locate the boundaries of syntactical units, which are also word boundaries. At these boundaries, a prosodic score can be joined by N-best rescoring to the acoustic and linguistic scores available in the speech recognizers. According to our experiments, prosody exploited in this way improves speech recognition performance (and can help the place punctuation marks as well).

We think that the developed prosodic segmenter can also be of interest in natural language processing tools, like syntactic analyzers.

Authors

GYÖRGY SZASZÁK graduated at the Budapest University of Technology and Economics in 2002. In the same year, he became research assistant at the Laboratory of Speech Acoustics, where his main research topics are speech recognition, pronounciation variation, speech database construction and the use of prosody in speech recognition. He is co-author of a dozen of articles and book chapters, mainly in the domain of prosody in speech recognition.

KLÁRA VICSI is the head of the Laboratory of Speech Acoustics at BME-TMIT. She became Doctor of Philosophy in 1992, Doctor of the Engineering Sciences of the Hungarian Academy of Sciences in 2004. She habilitated at BUTE in 2007. She was leader of several Hungarian and international research projects and she is also currently active project leader in the fields of speech acoustics, psychoacoustics, speech recognition and speech databases. She also participates in the development of speech aid systems for hard of hearing children or adults. She has more than 65 Hungarian and international publications and she is author of several book chapters on speech recognition.

References

[1] Gallwitz, F., Niemann, H., Nöth, E., Warnke, V.:
Integrated recognition of
words and prosodic phrase boundaries.
Speech Communication, Vol. 36, 2002,
pp.81–95.

[2] Kassai, Ilona:
Fonetika.
Tankönyvkiadó, Budapest, 1998.

[3] Kompe, R.:
Prosody in Speech Understanding Systems.
LNAI 1307, Springer Verlag, Berlin-Heidelberg, 1997.

[4] Kompe, R., Kiessling, A., Niemann, H., Nöth, H.,
Schukat-Talamazzini E. G., Zottman, A., Batliner, A.:
Prosodic scoring of word hypothesis graphs.
Proc. of the European Conference on Speech
Communication and Technology, Madrid, 1995.
pp.1333–1336.

[5] Riley, M., Byrne, W., Finke, M.,
Khudanpur, S., Ljolje, A.:
Stochastic pronunciation modelling from hand-
labelled phonetic corpora.
In: Modeling Pronunciation Variation for ASR, 1998.
pp.109–116.

[6] Roach, P. S. et al.:
BABEL:
An Eastern European Multi-language database.
International Conf. on Speech and Language, 1996.

[7] Sjölander, K. and Beskow, J.:
Wavesurfer – an open source speech tool.
Proceedings of the 6th International Conference of
Spoken Language Processing,
Beijing, China, 2000. Vol. 4,
pp.464–467.

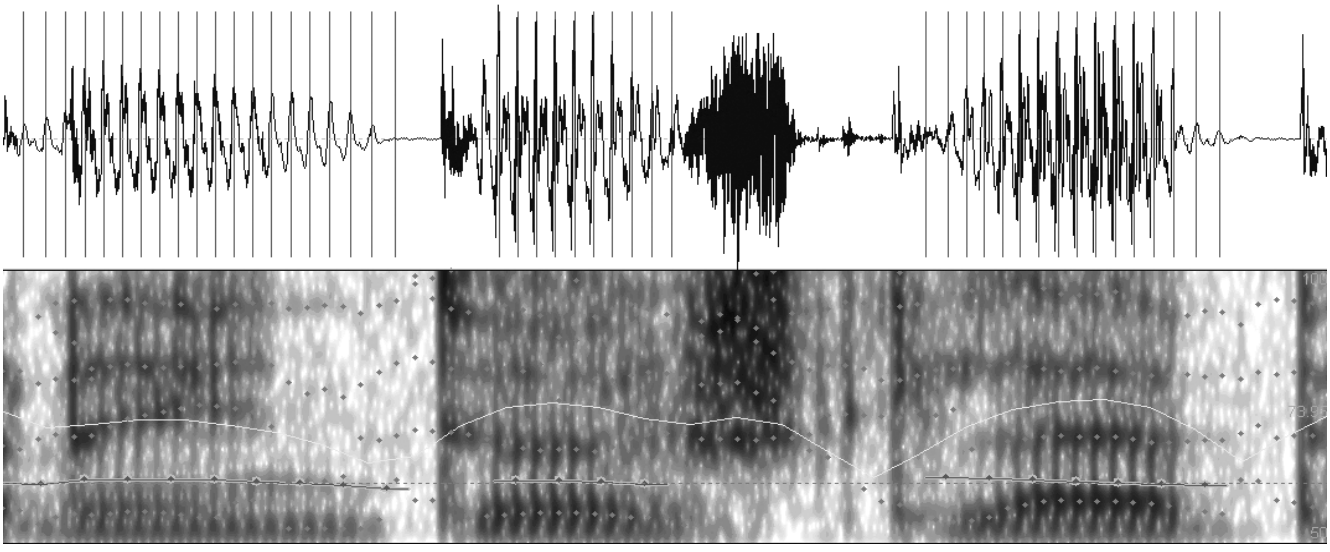
[8] Szaszák, Gy. - Vicsi, K.:
Folyamatos beszéd szószintű szegmentálása
szupra-szegmentális jegyek alapján.
In: III. Magyar Számítógépes Nyelvészeti Konf.,
Szeged, 2005.,
pp.360–370.

[9] Young, S. et al.:
The HTK Book (for version 3.3).
Cambridge University, 2005.

[10] Veilleux, N. M., Ostendorf, M.:
Prosody/parse scoring and its aopplication in ATIS.
In: Human Language and Language and Technology
Proc. of the ARPA workshop, Plainsboro, 1993.
pp.335–340.

[11] Vicsi K., Kocsor A., Tóth L., Velkei Sz., Szaszák Gy.,
Teleki Cs., Bánhalmi A., Paczolay D.:
A Magyar Referencia Beszédadatbázis és alka-
lmazása orvosi diktálórendszerek kifejlesztéséhez.
In: III. Magyar Számítógépes Nyelvészeti Konf.,
Szeged, 2005.,
pp.435–438.

[12] Vicsi, K., Szaszák, Gy.:
Automatic Segmentation of Continuous Speech on
Word level Based on Supra-segmental features.
In: International Journal of Speech Technology,
Vol. 8, No.4, 2005.,
pp.363–370.



Speech enhancement in the reconstructed phase-space

ISTVÁN PINTÉR

Kecskemét College, GAMF Faculty, Sándor Kalmár Institute of Information Technology
pinter.istvan@gamf.kefo.hu

Keywords: speech enhancement, signal subspace, reconstructed phase space, dimension embedding

The speech enhancement method, presented in this paper, is based on the concepts of reconstructed phase-space and dimension embedding. The proposed algorithm separates the speech from noise using a non-linear transformation in a transformed domain. Our recent results in case of uncorrelated, additive noise are presented in this paper.

1. Introduction

Speech enhancement is a long-standing problem in digital speech processing [1]. Several methods for noise suppression have been elaborated during the past three decades. The common assumption in most cases is the slow variation of noise parameters, corresponding to the linear speech model.

As an example, a system worth mentioning uses an auditory-model based filterbank with Wiener-filtering in sub-bands [2]. According to published results these methods give acceptable solutions in case of SNRs (signal to noise ratio) greater than 6...9 dB [3]. In case of either lower SNRs or nonstationary noise the speech enhancement methods are based on non-linear models. A non-linear model of human auditory system has been applied for noise suppression in [4], while the reconstructed phase-space representation of speech belongs to the class of non-linear signal models [5]. The latter is also the subject of the recent paper.

The structure of the paper is as follows. In the first part the optimal representation of the clean speech is reviewed, followed by the introduction of a noise suppression method based on the notion of speech subspace. The generalised version, working in the reconstructed phase space, is also introduced. Our numerical results, achieved by realisation of the algorithm are presented in the fourth section. The paper ends with the conclusions, acknowledgement and references.

2. Representation of the clean speech in the transformed domain and in the reconstructed phase space

The noise suppresion method, presented in this paper, is based on two assumptions. The first one is the existence of the optimal representation of the clean speech, the second one is that the concept of reconstructed phase space is suitable for speech processing problems.

Concerning the first assumption, a vector can be formed from α_n speech samples of the segment under pressing. If N denotes the number of samples in the segment, the resulting vector corresponds to a vector of N dimensional Euclidean-space. This vector $\underline{s}$ can be written as a linear combination using the $\{\underline{t}_n\}$ natural orthonormal basis, where coefficients are the speech samples: $\alpha_n = (s,\underline{t}_n)$ and the *n*th component of the N-dimensional $\underline{t}_n$ column vector is 1, the others are 0. According to experiences in solutions of practical problems in digital speech processing, there exists an orthonormal basis, so that by using this 'optimal' basis the speech vector can be represented with fewer components than N [6]. The optimality means that the speech vector in question can be given as

$$\hat{\underline{s}} = \sum_{n=0}^{L-1} a_n \cdot \underline{v}_n \tag{1}$$

where $\{\underline{v}_n\}$ denotes the optimal orthonormal basis and L<N holds. Moreover, the representation in (1) is optimal in the sense that the value of the criterion function below is

$$J(\underline{e}) = E\left\{\|\underline{e}\|^2\right\} = E\left\{\|\underline{s} - \hat{\underline{s}}\|^2\right\} = \sum_{n=L}^{N-1} \underline{v}_n^T \cdot E\left\{\underline{s} \cdot \underline{s}^T\right\} \cdot \underline{v}_n = \sum_{n=L}^{N-1} \underline{v}_n^T \cdot \underline{\underline{R}} \cdot \underline{v}_n \tag{2}$$

that is the mean square error is minimal (ideally L<N and $||\underline{e}||=0$). By the assumtion of $E\{\underline{s}\}=0$, we get $\underline{\underline{R}}=\underline{\underline{K}}$, which is the covariance matrix. The solution of (1) and (2) is the $\{\underline{v}_n\}$ eigenvector system of the covariance matrix, and the minimal value othe mean square error can be written using the corresponding eigenvalues as

$$J(\underline{e}) = \sum_{n=L}^{N-1} \lambda_n$$

where λ_n denotes the *n*th eigenvalue. The new representation of the speech vector $\underline{s}$ can be computed as a matrix-vector product using the matrix below

$$\underline{\underline{T}} = \left(\underline{v}_0^T, \underline{v}_1^T, \ldots, \underline{v}_i^T, \ldots, \underline{v}_{N-1}^T\right)^T \tag{3}$$

which has the eigenvectors in its rows corresponding to eigenvalues organized in desending order.

The second assumption goes for the representation of the speech in the reconstructed phase space. The concept of reconstructed phase space applies to the

motion equation of the discrete dynamical system, $\underline{x}_{n+1}=\underline{F}(\underline{x}_n)$, where $\underline{x}_n$ and $\underline{x}_{n+1}$ are D dimensional points in the phase space, $\underline{F}$ denotes a suitable mapping. The set $\{\underline{x}_n\}$ of phase-space points constitutes the so-called trajectory. This trajectory cannot be observed directly, only through the non-linear mapping $\underline{x}_n \rightarrow g(\underline{x}_n)$ – the resulting observable real number is the speech sample $\alpha_n = g(\underline{x}_n)$. By taking these samples in regular time intervals T_{MV}, one finally gets the speech sample sequence $\{\alpha_n\}$. It is provable that when the condition $M>2\cdot D+1$ holds, then from the number sequence α_n the vector sequence $\{\underline{y}_n\}$ can be reconstructed, which is equivalent of the original vector sequence $\{\underline{x}_n\}$. The method of the reconstruction is the so-called dimension embedding, which results a vector

$$\underline{y}_n(M,\tau)=(\alpha_n,\alpha_{n+\tau},\ldots,\alpha_{n+(M-1)\cdot\tau}) \tag{4}$$

where $\tau>0$ is the time lag (given by number of samples here), and $M>0$ denotes the embedded dimension. The equivalence mentioned above means that there exists an invertible, smooth mapping $\underline{h}:\underline{y}_n\,(M,\tau)\rightarrow\underline{x}_n$, by which the two vector sequence in question can be transformed into each other [7]. The values of the embedding dimension M and time lag τ can be determined experimentally, depending on the type of the speech technology application. According to relevant literature the value of the embedding window $M\cdot\tau\cdot T_{MV}$ is in the interval of 1...5 ms [8].

3. Noise suppression in the reconstructed phase space by using the sub-space method

The noise suppression algorithm, which can be given by using the concept of the reconstructed phase space, is in essence a generalisation of an earlier method published in the relevant literature, so the latter is reviewed first.

The basis of the method is the property of the speech described in Section 1, namely that the speech can be optimally represented. It means, that the N dimensional orthonormal basis is not necessary for the representation, but L<N dimensional orthonormal basis is enough, and ideally the mean square error value is zero. So, the N dimensional speech vector can be found in an L dimensional sub-space, titled as 'speech-subspace'.

The noise suppression algorithm determines an estimated, optimal speech vector from the noisy speech samples. Let's denote the noisy speech as

$$\underline{u}=\underline{s}+\underline{w} \tag{5}$$

where $\underline{w}$ denotes the additive noise vector, uncorrelated with speech. Starting from the noisy samples, it is necessary to give an estimate of the speech $\tilde{\underline{s}}$, so that the expectation value of the norm of the diference $\underline{s}-\tilde{\underline{s}}$ should be minimum, that is

$$E\left\{\left\|\underline{s}-\tilde{\underline{s}}\right\|^2\right\}\rightarrow\min \tag{6}$$

First of all – similarly to the above discussed problem – it is necessary to determine the optimal orthonormal basis for the speech, however in this case only the noisy speech vector $\underline{u}$ is known. By assuming that $E\{\underline{w}\}=0$, and using the previous assumption $E\{\underline{s}\}=0$, gives $E\{\underline{u}\}=0$. Additional assumption is that the zero-mean noise is white noise, if its covariance-matrix can be written as $\underline{\underline{K}}^{NOISE}=\sigma^2\cdot\underline{\underline{I}}$, where $\sigma>0$ and $\underline{\underline{I}}$ denotes the NxN identity matrix. Because the speech and noise are uncorrelated, the correlation matrix of the noisy speech can be written as a sum of correlation matrices of speech and noise, respectively, that is

$$\underline{\underline{K}}^{NOISY}=E\left\{\underline{u}\cdot\underline{u}^T\right\}=\underline{\underline{K}}^{SPEECH}+\underline{\underline{K}}^{NOISE} \tag{7}$$

holds. As it can also be proven, the eigenvectors of the noisy and clean speech are the same. The latter property makes it possible to determine the estimated speech vector, because the vectors of the orthonormal basis, necessary for the ideal representation of the speech, can be determined from the given noisy speech samples. In other words, the optimal basis $\{\underline{v}_n\}$ can be computed from the covariance matrix of the noisy speech, so it is not necessary to know the covariance matrix of the clean speech. Moreover, as a consequence of the summability of the covariance matrices (7), it is provable, that the covariance matrix of the noisy speech in the transformed domain is the diagonal matrix below:

$$\underline{\underline{K}}^{NOISY}\Big|_{\{\underline{v}_n\}}=diag\left(\lambda_0+\sigma^2\ \ldots\ \lambda_{L-1}+\sigma^2\ \ \sigma^2\ \ldots\ \sigma^2\right). \tag{8}$$

According to our assumption described in Section 1, the speech can optimally be represented in the subspace, spanned by the vectors $\underline{v}_0,\underline{v}_1,\ldots,\underline{v}_{L-1}$. In other words, in case of noisy speech, in this sub-space both speech and noise 'can be found', while in the orthogonal complement, that is in the sub-space, spanned by the vectors $\underline{v}_L,\underline{v}_{L+1},\ldots,\underline{v}_{N-1}$, only noise 'can be found'.

The noise suppression algorithm should be given in the form of a linear transformation $\underline{\underline{H}}$, that is

$$\tilde{\underline{s}}=\underline{\underline{H}}\cdot\underline{u}\,. \tag{9}$$

The estimation error is the remainder $\underline{r}=\underline{s}-\tilde{\underline{s}}$. The authors of [6] demonstrated, that the remainder signal

$$\underline{r}=\underline{r}^{SPEECH}-\underline{r}^{NOISE} \tag{10}$$

has two components. One of them correlated with the speech, while the other is correlated with the noise. Because of this the task is not only to minimize the speech-correlated component, but to suppress the noise-correlated component in a prescribed manner. This problem has been solved in [6] both in time domain and in spectral domain. Our results concerning the time domain have been published in [9]. In spectral domain it is also necessary to minimize the speech-correlated component, however it is possible to specify a noise suppression condition for every spectral component. Thus, the noise suppression problem can be formulated as a constrained optimization problem:

$$J\left(\underline{r}^{SPEECH}\right)\xrightarrow[\underline{\underline{H}}]{}\min \tag{11}$$

so that:

$$E\left\{\left|\underline{v}_n^T \cdot \underline{r}^{NOISE}\right|^2 \le \beta_n \cdot \sigma^2\right\} \quad n = 0,1,...L-1, \text{ and}$$
$$E\left\{\left|\underline{v}_n^T \cdot \underline{r}^{NOISE}\right|^2 = 0\right\} \quad n = L,...N-1. \tag{12}$$

The first condition is a component-wise specification for the remainder noise in the speech sub-space with conditions of $\beta_n > 0$, respectively, while the second condition is simply the zeroing the components in the noise sub-space. The optimal transformation matrix can be given by using the Karush-Kuhn-Tucker constraint optimization method as [6]:

$$\underline{\underline{H}}^{OPT} = \underline{\underline{V}} \cdot \underline{\underline{G}} \cdot \underline{\underline{V}}^T,$$
$$\underline{\underline{G}} = \mathrm{diag}\left(g_{0,0},..., g_{N,N}\right),$$
$$g_{n,n} = \begin{cases} \sqrt{\gamma_n} & n = 0,1,...,L-1 \\ 0 & n = L,...,N-1 \end{cases} \tag{13}$$

where the column vectors of the matrix $\underline{\underline{V}}$ are the eigenvectors. For the values of γ_n two methods can be found in [6]. In this paper the relationship below

$$\gamma_n = \exp\left(-\frac{\kappa \cdot \sigma^2}{\lambda_n^{SPEECH}}\right) \tag{14}$$

has been used. The degree of noise suppression can be set up with the experimental constant of $\kappa \ge 1$, also affecting the distorsion of the estimated speech.

The method described above can be generalised to the case of reconstructed phase space. Namely, the latter as a model background makes it possible to generate an M-dimensional data set from a given single noisy vector $\underline{u} = \underline{s} + \underline{w}$ by using the method of dimension embedding. Because of its construction, for the resulting trajectory matrix $\underline{\underline{U}}_{MxN}$ the following relationship holds

$$\underline{\underline{U}}_{MxN} = \underline{\underline{S}}_{MxN} + \underline{\underline{W}}_{MxN}. \tag{15}$$

Moreover, because for every corresponding sample the relationship $u_n = s_n + w_n$ holds, for the trajectory matrix-based covariance matrix we obtain:

$$\underline{\underline{K}}_{\underline{\underline{U}}_{MxN}} = \underline{\underline{K}}_{\underline{\underline{S}}_{MxN}} + \underline{\underline{K}}_{\underline{\underline{W}}_{MxN}}, \tag{16}$$

where

$$\underline{\underline{K}}_{\underline{\underline{W}}_{MxN}} = \sigma^2 \cdot \underline{\underline{I}}_{MxM}$$

Because of this, the noise suppression procedure above can also be applied for the estimation of the trajectory matrix $\underline{\underline{\tilde{S}}}$. Finally, from a given trajectory matrix estimate it is necessary to determine a speech vector estimate $\underline{\tilde{s}}$, which can be peformed based on the construction of the matrix $\underline{\underline{U}}$. This latter method differs from the original sub-space method not only in determining the data set necessary to determine the covariance matrix, but in the estimation of the speech sample as well, because the phase space-based method results several speech sample estimates for a given sample.

The trajectory matrix used in our work is based on a periodic extension of the noisy speech segment, so every speech sample has exactly M estimates, and the final estimate is their average. That is, in our case the weighting matrix for the final estimate is not necessary, while in other constructs it is needed [10]. More formally, the element $u_{i,j}$ of our trajectory matrix can be given as

$$u_{i,j} = u_{(j+i\cdot\tau) \bmod N}, \tag{17}$$

where N denotes the number of segment's samples, M denotes the embedding dimension, and τ denotes the time-lag. It is worthwile mentioning that our covariance matrix also differs from the empirical Toeplitz covariance matrix of [6] and from those published in [5] and [10].

4. Realisation of the speech enhancement algorithm and numerical results

In this work our goal was to demonstrate the method and the algorithm, so we have analysed only one long Hungarian sentence. The sentence was uttered by a native Hungarian male speaker, and the speech has been sampled with 8 kHz sampling frequency followed by a 16 bit linear quantisation. The resulting speech sample sequence was the 'clean' speech. However, even in this case the value of the global SNR was 45,8 dB (computed in active speech regions only). The noisy speech has been computed using these samples by artificially adding noise. The source of the noise samples is a part of the RSG-10 noise database [11]. Because of the different sampling frequencies, a suitable re-sampling was necessary before addition. The noise types, investigated in this work were the following: white noise, pink noise, high frequency channel noise. The noise level has been set up using the energy of the clean speech, computed in active speech regions.

The effectiveness of noise suppression has been charecterized by the number below

$$SRR = 10 \cdot \lg\left(E^{SPEECH} / E^{RESIDUAL}\right), \tag{18}$$

(Signal to Residual Ratio), where the nominator is the speech energy in the active speech region, and the denominator is the energy of the residual signal (computed on the same index-set as the nominator).

The noise suppression has been applied to a sequence of overlapped speech segments, using 50% overlap. The segment has been windowed using a Hanning-window before enhancement, and the final estimation has been computed by the overlap-add re-synthesis technique. The segment length, the embedding dimension, the time lag, the dimension of the speech sub-space and the value of the constant κ has been determined experimentally by many listening tests. The parameters γ_n, necessary for the spectral method, have been estimated as follows.

The value of σ^2 has been estimated by the first eigenvalue of the noise sub-space, while the values of λ_n^{SPEECH} have been estimated with the difference between the eigenvalues in the speech sub-space and the estimated value of σ^2. The computation of the eigenvalues and ei-

genvectors based on Jacobi's algorithm, and the noise suppression algorithm has been realised in C. *Table 1* contains the numerical results obtained – they correspond to those of published in the relevant literature [5,10].

SNR (dB)	SRR (dB)		
	White noise	High frequency channel noise	Pink noise
15	9,3	9,3	9,0
12	9,1	9,0	8,3
9	8,7	8,6	7,3
6	8,1	7,9	5,7
3	7,2	6,9	3,7
0	6,0	5,7	1,8
-3	4,5	4,2	0,0

Table 1.
The SRR values in cases of different SNRs and noise types (segment length: 800 sample, embedding dimension: 20, time-lag: 1 sample, speech sub-space dimension: 7, empirical constant: κ=5)

It is seen from Table 1, that SRR values are greater than SNRs, only if SNRs are lower than 6 dB. The reason is a property of the method itself discussed in Section 2, namely it is not only supresses the noise, but distortes the speech as well. The graphical illustration of the the algorithm can be seen in *Figure 1*.

It impressively demonstrates the noise suppression capability of the algorithm and also its speech distortion effect.

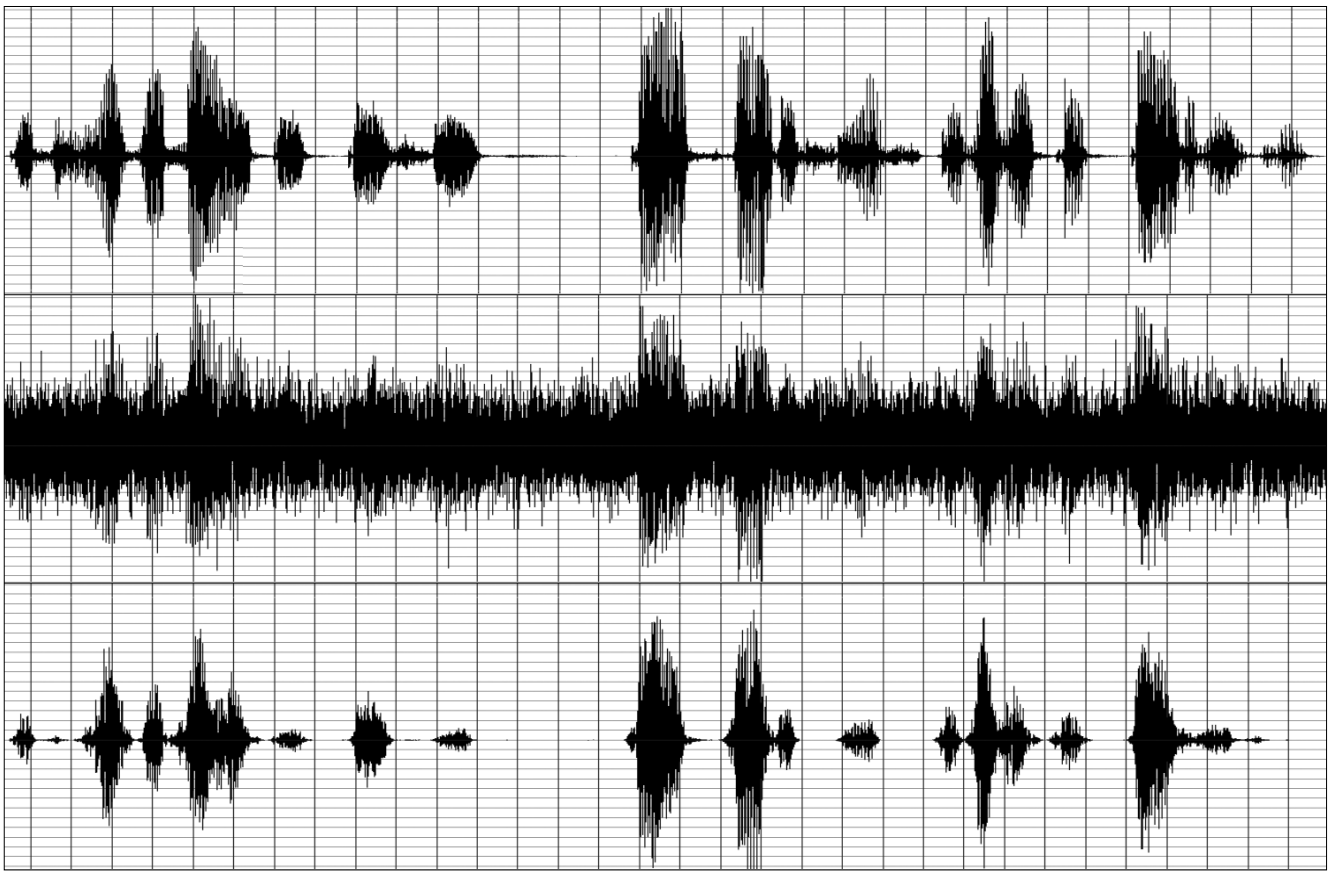
5. Conclusions

We discussed a speech enhancement algorithm, working in the reconstructed phase space.

The algorithm is based on dimension embedding, and assumes the separability the speech sub-space and the noise sub-space in the Euclidean space, determined by the covariance matrix of the data set after embedding. The Euclidean space in question is spanned by the eigenvectors of the covariance matrix above and the eigenvectors have been computed using Jacobi's algorithm. The data set has been determined after periodic extension of the speech segment, which differs from the published methods. That means, the so-called weighting matrix is not necessary for the estimation of the speech sample in our method.

The program has been tested using a Hungarian sentence by artificially added noise using three noise types and seven different noise levels. The enhancement capability has been determined numerically, the parameters have been set up experimentally by many listening tests. The best results have been achieved using the parameters as follows: about 100 ms segment length, 50% segment overlap, Hanning window, overlap-add re-

Figure 1.
Noise suppression in case of -3 dB and white noise in case of the same parameters as in Table 1. (upper trace: original utterance, middle trace: noise speech, lower trace: enhanced speech)



synthesis, 20 dimensional embedding space, 1 sample time lag, 7 dimensional speech sub-space.

The values correspond well to those published in the literature, not only in case of white noise, but in case of high frequency channel noise and pink noise. However, the algorithm is optimal only in case of white noise, for other noise types it is necessary to apply a whitening transformation.

Our further work is the automatic determination of the values of the embedding dimension, the time-lag and the dimension of speech sub-space, moreover the testing of the method using a large noisy speech database.

Acknowledgement

The author would like to thank Géza Gordos, Géza Németh and Péter Tatai for their kind help and encouragement in his speech processing algorithm development work.

Author

ISTVÁN PINTÉR received his MS degree in electrical engineering in 1983 and his PhD degree in informatics in 1997 from the Technical University of Budapest. In 1983 he joined the MIKI and in 1984 the GAMF (now College of Kecskemét, GAMF Faculty), where he is a professor. His research interests are digital speech processing (novel speech representations, speech enhancement), digital signal processing (discrete orthogonal transforms), pattern recognition (application of artificial neural networks). He has published several journal articles and conference papers in the areas above. He has received the Pollák-Virág award from HTE in 2007.

References

[1] J. S. Lim, A. V. Oppenheim:
Enhancement and bandwidth compression of noisy speech.
Proceedings of IEEE 67 (12), 1979, pp.1586–1604.

[2] Yang Gui, Kwan, H. K.:
Adaptive sub-band Wiener filtering for speech enhancement using critical-band gammatone filterbank,
Proceedings of 48th Midwest Symposium on Circuits and Systems, 2005, Vol. 1, pp.732–735.

[3] Haci Tasmaz, Ergun Ercelebi:
Speech enhancement based on undecimated wavelet packet-perceptual filterbanks and MMSE-STSA estimation in various noise environments.
Digital Signal Processing,
(p.16, in press, available online 12 October 2007).

[4] T. F. Quatieri, R. B. Dunn:
Speech enhancement based on auditory spectral change.
Proceedings of International Conference on Acoustics, Speech and Signal Processing, Orlando, Florida, IEEE, 13-17 May 2002, pp.257–260.

[5] J. Sun, N. Zheng, X. Wang:
Enhancement of Chinese speech based on nonlinear dynamics.
Signal Processing 87, 2007, pp.2431–2445.

[6] Y. Ephraim, H. L. Van Trees:
A Signal Subspace Approach for Speech Enhancement.
IEEE Trans. on Speech and Audio Processing, Vol. 3, No.4., July 1995, pp.251–266.

[7] H. Kantz, T. Schreiber:
Nonlinear Time Series Analysis.
Cambridge University Press, 1997.

[8] G. Kubin, C. Lainscsek, E. Rank:
Indentification of Nonlinear Oscillator Models for Speech Analysis and Synthesis.
In: Chollet et al. (eds.):
Nonlinear Speech Modeling.
LN AI 3445, Springer Verlag, 2005, pp.74–113.

[9] I. Pintér:
Noise suppression using non-linear speech model.
Pollack Periodica, Vol. 2, Supplement, Akadémiai Kiadó, 2007, pp.121–133.

[10] M. T. Johnson, R. T. Povinelli:
Generalized phase space projection for nonlinear noise reduction.
Physica D 201, 2005, pp.306–317.

[11] http://spib.rice.edu/spib/select_noise.html

NETWORKS 2008

13th International Telecommunications Network Strategy and Planning Symposium

„Convergence in Progress"

DANUBIUS HEALTH SPA RESORT MARGITSZIGET • BUDAPEST, HUNGARY • September 28 – October 2, 2008

CALL FOR PARTICIPATION

SCOPE

Networks 2008 will focus on the challenges of planning networks to deliver on the promise of convergence of the information and communication technologies (ICT) and next generation networks (NGN).
It is the 13th of such symposia, held every two years, and attracting participants from all over the world: network operators, software companies, system integrators, researchers from universities and industry, marketers and policy-makers, regulators.
At **Networks 2008** we continue the tradition of state-of-the art papers, invited presentations, and panel sessions, as international experts present their latest findings and share experiences in network strategy, planning, operations, management, control and design.

PROGRAM (PRELIMINARY)

Networks 2008 will feature sections on topics such as:

- **Broadband Access: Technology and Economics**
- **Service & Technology Drivers for Optical Networks**
- **Scalable Bridging, Routing & Admission, Resilience Schemes**
- **Routing over WDM Optical Networks**
- **Multi-Layer Networks: Routing, Traffic Eng., Modeling, Design**
- **Multi-Domain Networks: BGP & PCE**
- **Wireless Networks: Optimizing Scarce Resources**
- **Mobile Networks**
- **Optimizing Network Operation**
- **Traffic Analysis and Modeling**
- **Network Security**
- **Migration towards FMC: Control & Addressing**
- **Towards IMS: Video Services in Focus**
- **NGN Services: Convergence, Migration, Market, Business**

TUTORIALS

Networks 2008 will feature sections on topics such as:

- **From opacity to transparency via translucent optical networks**, Maurice Gagnaire, *ENST Paris, France*
- **Efficiency of P2P networks and alternative overlays for content delivery and internet service provisioning**, G. Hasslinger, *T-Systems*
- **IPv6 training**, János Mohácsi, *NIIF, Hungary*
- **Fiber-to-the-X: Technologies & Economics**, Mohamed El-Sayed, *Alcatel-Lucent, USA*
- **Focusing the study on wireless multihop networks**, Josep Paradells, Carles Gomez, *Technical University of Catalonia, Spain*
- **End to end IPTV Design and Implementation – How to avoid pitfalls?**, Siri Hewa, *Ericsson, Sweden*
- **Carrier Ethernet Transport in Metro and Core Networks**, Claus G. Gruber and Achim Autenrieth, *Nokia Siemens Networks*
- **IP-Oriented QoS in the Next Generation Networks: application to wireless networks**, Pascal Lorenz, *University of Haute Alsace, France*
- **Next Generation Network Techno-Economic Insights**, Santiago Andrés Azcoitia, *Telefónica I+D, Spain*
- **Changing paradigms from bridged to carrier Ethernet**, Dimitri Papadimitriou, *Alcatel-Lucent Bell, Belgium*
- **NGN architectures and their management**, Idir Fodil, *FT R&D, France*
- **Practical steps in techno-economic evaluation of network deployment planning**, Sofie Verbrugge, *UGent/IBBT–INTEC, Belgium*

KEYNOTE TALKS

Besides the technical sessions, held in three parallel tracks, the conference will also include several keynote talks and panels, featuring highly recognized specialists from regulatory bodies, industry and academia.

Opening Session:
- Opening words: Peter Kiss, Minister of Prime Minister's Office
- Challenges in ICT regulation: Representative of IRG/ERG
- Christopher Mattheisen, Chairman & CEO, Magyar Telekom: Vision and objectives of Hungarian telecommunications

Plenary – Convergence in Progress, Two Perspectives:
- Cayetano Carbajo Martín, Telefonica: Strategy, evolution, fixed network, applications and services, market in Latin-America
- Rati Thanawala, Alcatel-Lucent, USA: A global perspective

Plenary - Wireless Communication Panel:
- chaired by Alberto Ciarniello, Telekom Italia

Closing Session:
- Bernard Jarry-Lacombe, France Telecom: Access and fix networks
- Andy Valdar, UCL, UK: What's driving the adoption of next generation network (NGN) and next generation access (NGA)

FREE TUTORIALS

The registration for tutorials is FREE for students, and for IEEE/HTE members registering for the conference.

IMPORTANT DATES

Early bird registration deadline: AUGUST 1, 2008
On-line registration deadline: September 21, 2008
Conference dates: September 28 – October 2, 2008

COMMITTEES

International Management and Scientific Committee (IMSC)
IMSC Chair: Gyula Sallai, *BME-TMIT and HTE, Hungary (sallai@tmit.bme.hu)*
IMSC Members: Alberto Ciarniello, *Telecom Italia Mobile, Italy*
Oscar Gonzalez-Soto, *ITU Consultant, Spain*
Joachim Gross, *Arcor, Germany*
Wolfgang Gross, *Telekom, Germany*
Hideaki Yoshino, *NTT, Japan*
Bernard Jarry-Lacombe, *France Telecom, France*
Sang-Baeg Kim, *Korea Telecom, Korea*
Hussein T. Mouftah, *University of Ottawa, Canada*
Lawrence Paratz, *Telstra Corporation, Australia*
Michal Pióro, *Warsaw University of Technology, Poland*
Rati C. Thanawala, *Alcatel-Lucent, USA*
Andy Valdar, *University College London, UK*

International Programme Committee (IPC)
Chair: Tibor Cinkler, *BME-TMIT, Hungary (cinkler@tmit.bme.hu)*

National Advisory Board (NAB)
Chair: László Pap, *BME-HIT and HTE, Hungary (pap@hit.bme.hu)*

Local Organizing Committee (LOC)
Chair: Péter Nagy, *HTE, Hungary (nagy.peter@hte.hu)*

NETWORK 2008 is organized by:

Free Tutorial for Students | Early Deadline for Registration: August 1, 2008

www.networks2008.org

Patrons:

Technical co-sponsors:

OSNR based routing in WDM optical networks

SZILÁRD ZSIGMOND, MARCELL PERÉNYI, TIBOR CINKLER

Department of Telecommunication and Media Informatics, Budapest University of Technology and Economics

{zsigmond; perenyim; cinkler}@tmit.bme.hu

Keywords: routing, signal power, WDM, cross-layer optimization OSNR, RWA

In this paper we propose new physical impairments based routing and wavelength assignment (ICBR) methods where the control plane has influence on the signal power of the Wavelength Division Multiplexed (WDM) channels in metro-optical networks. We give the exact integer linear programming (ILP) formulation of the method. Nowadays in nearly all kinds of reconfigurable optical add-drop multiplexers (ROADM) the signal power can be tuned via variable optical attenuators (VOA) from the control plane. The proposed algorithm can be used in existing WDM optical networks where the nodes support signal power tuning. The proposed method outperforms the traditional existing schemes.

1. Introduction

WDM networks have successfully solved the capacity issues, but the continuously changing traffic still causes a serious problem for the operators. Emerging demands often cannot be satisfied without modifying the network design, which is quite costly and difficult, so operators try to avoid this situation whenever possible. There is a strong need for a system that can deliver the same capacity as WDM, with the design and provisioning flexibility of SONET/SDH. The solution must ensure flexibility for dynamically changing future demands.

The reconfigurable optical network offers the possibility to increase or change services between sites with no advanced engineering or planning, and without disrupting existing services. In the past, reconfigurable optical networking technology was too expensive or delicate to be widely deployed. With recently matured silicon-based integrated Planar Lightwave Circuit components, reconfigurable optical add/drop multiplexers (ROADMs) are now being installed by many operators. The technology called ROADM represents a real breakthrough for WDM networks by providing the flexibility and functionality required in present complex networking environments. Older, or fixed, OADMs cannot configure capacity at a node without manual reconfiguration and typically support reconfiguration of only a limited number of wavelengths. In contrast, ROADMs allow service providers to reconfigure add and drop capacity at a node remotely, reducing operating expenses by eliminating the time and complexity involved in manual reconfiguration.

ROADM by itself is not enough. Increased data management capabilities on individual wavelengths are also needed to exploit the benefits of ROADM in metro and backbone WDM networks. For instance, ROADM rings are very sensitive to topology changes and need strict monitoring and control of wavelength power to keep the system in balance. The real innovation lies in the system engineering related to the ROADM function, addressing per-wavelength power measurement and management, and per-wavelength fault isolation. Almost every optical system vendor has commercial ROADM with wavelength monitoring functions (see e.g. [1-4]).

The next step towards a fully reconfigurable WDM optical network is the deployment of tunable Small Form-factor Pluggable (SFP) interfaces, where the wavelength allocation is modified according to the network changes. The tunable dispersion compensation elements mean another innovation. Nowadays these ready-made products can be purchased [5,6]. The evolution of optical networks seems to tend towards a fully reconfigurable network where the control and the management plane (CP and MP) have new functions, such as determining the signal quality, tuning the wavelength frequency, setting dispersion compensation units, and – by using variable optical attenuators – setting the channel powers. Of course traditional functions (such as routing) remain the main function of the CP and MP. Routing and Wavelength Assignment (RWA) play a central role in the control and management of optical networks. Many excellent papers deal with the design, configuration, and optimization of WDM networks (see e.g. [7-8]). However, they do not consider the physical parameters of the fully reconfigurable optical network in the RWA method at all.

In this paper we propose a new ILP based RWA algorithm where the control plane handles the routing and the signal power allocation jointly. Nowadays in nearly all types of ROADMs the signal power can be tuned with variable optical attenuators (VOA) from the management system.

In metro WDM networks the signal power of the optical channels is determined by Cross-Phase (XPM) modulation and Raman scattering and not by the Brillouin threshold. This means that the total power inserted in fiber and not the channel powers has an upper bound. In this case it is possible to increase the powers of some channels up to the Brillouin threshold and at the same time the other channel powers have to be decreased to

fulfill the XPM and Raman scattering constraints. The previously mentioned idea can be used while configuring lightpaths. Let us assume a very simple scenario, see *Fig. 1* In this case we have two wavelengths λ_1 and λ_2. In Case "A" due to physical constraints node A can only reach node C in all-optical way. If there is a demand between node A-D this can only be established with signal regeneration or in node B or in C. In case "B" it is possible to increase the signal power of λ_2 to fulfill the OSNR request at the node D. in this way it is possible to established an all-optical connection between nodes A-D.

The proposed method can be used in existing WDM optical networks where the nodes support signal power tuning. The method also finds global optimum if it exists.

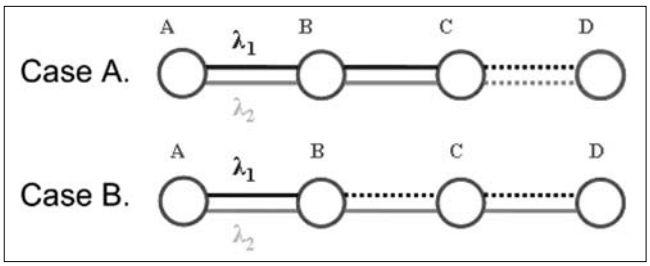


Figure 1.
The difference of OSNR based routing and traditional routing schemes

2. Physical feasibility

As mentioned before, our proposed algorithms use different channel powers in the same optical fiber. This approach introduces many new problems related to physical feasibility. All physical effects were already investigated using equal channel power allocation. The only difference in our case is that the impacts of the effects are different for each channel, since the signal powers are different. In case of linear effects the signal power has no influence on the dispersion and its compensation schemes. The only linear effect which has signal power dependency is the crosstalk in the nodes. We assume that using the well-known power budget design process the effects of the crosstalk can be eliminated.

More interesting question is how the EDFAs react to the use of different channel power allocations. For this purpose we made simulations using the VPI TMM/CM Version 7.5 simulation tool [9]. We assumed a system with 8 channels which are multiplexed and then amplified using EDFA rate propagation modules. We aimed at investigating the difference between the uniform and the adaptive channel power allocations. These results lead us to the conclusion that the so far deployed EDFAs behave similarly in case of both uniform and non-uniform channel power allocations in a single optical fiber.

The other interesting question is about the nonlinear effects, since these effects highly depend on the used signal powers. The only solution is to limit the signal power inserted in one optical fiber. This must be done in both allocation schemes. Another problem is the maximum allowed difference between the maximum and the minimum channel powers. In our case this is an input parameter of the algorithm. Determining this value is a hard task and is out of the scope of this paper. Finally to conclude: according to the results adaptive signal power allocation scheme can be implemented in optical systems deployed so far. Moreover, the authors know existing WDM optical networks operating without any error, where different channel powers – though not intentionally – are used, since the power tuning was not performed for the inserted channels in the ROADMs.

To investigate the relation between the signal power and the maximum allowed distance, we consider a noise limited system where other physical effects can be taken into account as power-penalty. It is possible to prove by analytical calculations that there is a linear relationship between the channel power and the maximum allowable distance of an all-optical link [10,11]:

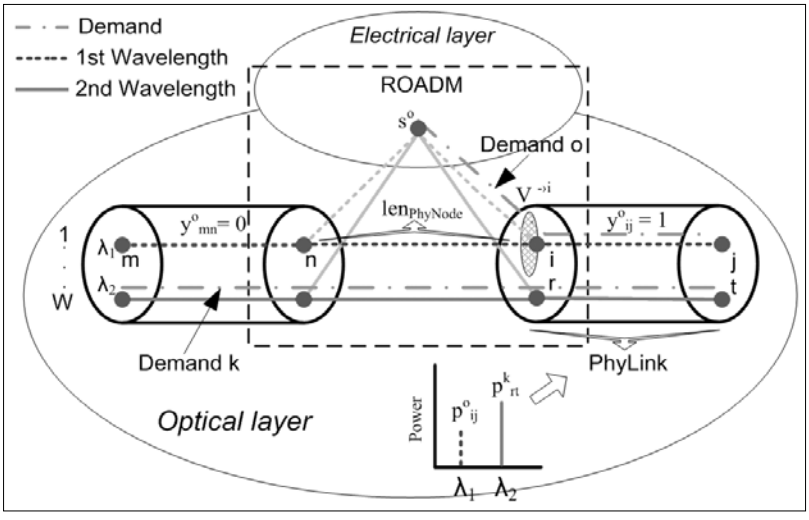
$$L = L_c \cdot P_{mW} \qquad (2.1)$$

where P_{mW} is the input power in mW, L is the maximum allowable distance, and L_c is the linear factor between it. For typical constant values, used in telecommunications, L_c is between 500 and 2000.

3. Network and Routing Model

We applied the wavelength graph (WL graph) modeling technique. The WL graph (which can be regarded as a detailed virtual representation of the network) is derived from the physical network considering the topology and the switching capabilities of the devices (nodes). The technique allows arbitrary mesh topologies, different types of nodes and joint optimization of multiple layers. A simpler version of the model has been first proposed in [12].

Figure 2.
Model of switching device with optical and electronic switching capabilities, grooming and 3R regeneration (in the electronic layer)



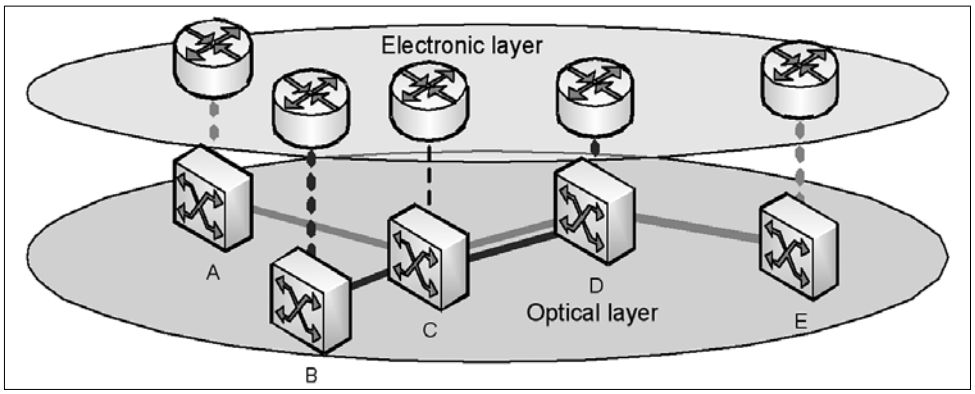


Figure 3.
End-to-end lightpaths are assigned to each demand. In this example there are two demands (A-E and B-D). Two lightpaths (A-E, B-D) are allocated, no grooming is allowed on link C-D.

The model of an ROADM switching device assumed in our simulations is shown in *Fig. 2*. The device can perform optical switching and – through the electronic layer – wavelength conversion, grooming and 3R signal regeneration. The device is illustrated in Figure 2 has an input and an output interface with a *physical link* (or fiber) connected to each. Each physical interface supports two *wavelengths* (*W*=2), marked by dark dashed and solid lines. The signal powers of the wavelengths in the right hand side physical link are different – as shown by the small subfigure. The example also comprises two demands (indicated by dash-dotted line): *demand k* passes through the switch in the optics, while *demand o* originates in this device in the electronic layer (s^o). A certain length of fiber ($len_{PhyNode}$) is assigned to each internal edge, e.g., edge *(n, i)*, which corresponds to the amount of signal distortion that the switching functionality introduces in the demand path. The edges representing O/E or E/O conversion are marked by grey color.

In routing we assume that WL conversion, grooming and signal regeneration are possible only in the electronic layer, and that the *noise and the signal distortion accumulate along the lightpath*. Actually, re-amplification, re-shaping and re-timing – which are collectively known as 3R regeneration – are necessary to overcome these impairments. Although 3R optical regeneration has been demonstrated in laboratories, only electronic 3R regeneration is economically viable in current networks.

The constraints of maximum input power in each fiber, and maximum allowed distance as a function of the input power of the lightpath have to be met.

In addition we differentiate between two routing cases and propose an ILP formulation for each (presented in Sections 4 and 5).

In the first case (referred to as *single-layer network*) we assume that a whole lightpath is assigned to each demand from source to destination node. The signal enters into the optical layer at the source node and leaves it at the destination node. Wavelength conversion, grooming or regeneration is not allowed elsewhere along the path.

In the second case (referred to as *multilayer network*) the path of a demand may consist of several lightpaths, i.e. it can enter and leave the electronic layer multiple times if necessary and efficient. In addition, in the second case grooming is also applicable.

4. ILP formulation of OSNR based routing in single layer networks

In this section we introduce the ILP formulation of OSNR based routing for single-layer networks *(Fig. 3)*.

4.1 Constants

The WL graph contains nodes (*V*) and edges (*A*). Edge *(i, j)* represents one edge in the WL graph. $V^{\rightarrow i}$ and $V^{i\rightarrow}$ represent incoming and outgoing edges of node *i*, respectively. Symbol A^{sw} denotes the set of edges in the WL graph representing switching function inside a physical device; other edges represent wavelengths of a physical link (A^{pl}). The set of demands in the network is denoted by *O*.

$$P_{pl}^{max} = \text{4-20 dBm, typically 10 dBm} \quad (4.1)$$

Constant P_{pl}^{max} means the upper limit of total power in physical link *pl* in dBm. $P_{pl}^{max}{}_{lin}$ the same in mW.

$$len_{ij} \quad (4.2)$$

Constant len_{ij} is the length of the physical link which the wavelength belongs to.

$$len_{PhyNode} = \text{90 km, typically} \quad (4.3)$$

Constant $len_{PhyNode}$ corresponds to the length of the fiber a switching device induces to the path of the demand.

$$L_c = 1200 \quad (4.4)$$

Constant L_c is the factor of the linear relation between the input power of a demand (in mW) and the maximum distance the signal is allowed to reach.

$$\alpha \quad (4.5)$$

Constant α expresses tradeoff between optimization objectives: minimal routing cost or minimal power.

$$s^o,\ t^o \quad (4.6)$$

Symbols s^o and t^o represent source and target of demand *o*.

$$\beta = \frac{n}{W} \cdot P_{pl}^{\max}{}_{lin} \quad (4.7)$$

Constant β is the maximum allowed signal power for one channel in mW. Here *n* is integer a real number between 1 and W, and W is the number of wavelengths in a fiber.

4.2 Variables

$$p^o \in \left[0..\frac{\beta}{P_{pl\ lin}^{max}}\right],\ \forall o \in O \tag{4.8}$$

Variable p^o denotes the input power of demand *o* divided by $P_{pl\ lin}^{max}$.

$$p_{ij}^o \in \left[0..\frac{\beta}{P_{pl\ lin}^{max}}\right],\ \forall (i,j) \in A,\ \forall o \in O \tag{4.9}$$

Variable p_{ij}^o means the power of demand *o* on edge *(i, j)* divided by constant $P_{pl\ lin}^{max}$.

$$y_{ij}^o \in \{0,\ 1\},\ \forall (i,j) \in A,\ \forall o \in O \tag{4.10}$$

Variable y_{ij}^o tells whether demand *o* uses edge *(i, j)* or not. (E.g., variable $y_{mn}^o = 0$ in Fig. 2, since demand *o* does not pass through edge *(m, n)*, which represents the first wavelength. On the other hand, variable $y_{ij}^o = 1$, because demand *o* does use edge *(i, j)*.)

4.3 Objective function

Minimize:

$$\alpha \cdot \sum_{\forall o \in O} \sum_{\forall (i,j) \in A/A_{sw}} y_{ij}^o + (1-\alpha) \cdot \sum_{\forall o \in O} p^o \tag{4.11}$$

The objective function expresses that the sum of the used edges should be minimized together with the sum of input powers of demands. If we want to minimize the total cost of the routing, constant cost factors should be assigned to each edge.

Constant α decides whether optimization emphasis is on minimal routing cost (α is close to 1) or on minimal input power (α is close to zero).

4.4 Constraints

$$\sum_{\forall o \in O} \sum_{\forall (i,j)^{pl}} p_{ij}^o \le 1,\ \forall pl \in PhyLinks \tag{4.12}$$

$$p_{ij}^o \le y_{ij}^o,\ \forall (i,j) \in A,\ \forall o \in O \tag{4.13}$$

$$\sum_{\forall j \in V^{\to i}} p_{ji}^o - \sum_{\forall k \in V^{i \to}} p_{ik}^o = \begin{cases} -p^o & \text{if } i = s^o \\ 0 & \text{if } i \notin \{s^o, t^o\},\ \forall i \in V,\ o \in O \\ +p^o & \text{if } i = t^o \end{cases} \tag{4.14}$$

$$\sum_{\forall j \in V^{\to i}} y_{ji}^o - \sum_{\forall k \in V^{i \to}} y_{ik}^o = \begin{cases} -1 & \text{if } i = s^o \\ 0 & \text{if } i \notin \{s^o, t^o\},\ \forall i \in V,\ o \in O \\ +1 & \text{if } i = t^o \end{cases} \tag{4.15}$$

$$\sum_{\forall o \in O} y_{ij}^o \le 1,\ \forall (i,j) \in A \tag{4.16}$$

$$\sum_{\forall (i,j) \in A^{sw}} y_{ij}^o \cdot len_{PhyNode} + \sum_{\forall (i,j) \in A^{pl}} y_{ij}^o \cdot len_{ij} \le \le L(p^o) = L_c \cdot p^o \cdot P_{pl\ lin}^{max},\ \forall o \in O \tag{4.17}$$

4.5 Explanation

Constraint (4.12) expresses that the sum power of demands traversing a physical link (fiber) cannot exceed the maximum allowed power of that link. Constraint (4.13) tells that if the power of demand *o* in edge *(i, j)* is larger than zero, than edge *(i, j)* is used by demand *o*. Constraints (4.14) and (4.15) express the flow-conservation constraint of the power and of the *y* decision variables, respectively, for every demand. Constraint (4.16) guarantees that a given edge can be used by only one demand. Constraint (4.17) ensures that the total length of demand *o* should be less than the distance allowed by the input power of demand *o*.

5. ILP formulation of OSNR based routing in single layer networks

In this section we introduce the ILP formulation of Signal Power based Routing for multilayer networks, which can provide optimal solution for the joint problem of RWA with grooming and of determining the signal powers of lightpaths *(Fig. 4)*.

5.1 Variables and constants

The symbols are similar to those introduced in 4.1. In addition the set of lightpaths is denoted by *L*. A path in the WL graph is considered as a lightpath if it goes only in the optical layer without going up to the electronic layer. A lightpath does not traverse through any electronic node except for the source and destination nodes.

$$p_{EF} \in \left[0..\frac{\beta}{P_{pl\ lin}^{max}}\right],\ \forall (E,F) \in L \tag{5.1}$$

Variable p_{EF} denotes the input power of lightpath *(E, F)* divided by constant $P_{pl\ lin}^{max}$.

$$p_{ij}^{EF} \in \left[0..\frac{\beta}{P_{pl\ lin}^{max}}\right],\ \forall (i,j) \in A,\ (E,F) \in L \tag{5.2}$$

Variable p_{ij}^{EF} means the power of lightpath *(E, F)* on edge *(i, j)* divided by constant $P_{pl\ lin}^{max}$.

$$x_{EF\ ij}^o \in \{0,\ 1\},\ \forall (i,j) \in A,\ o \in O,\ (E,F) \in L \tag{5.3}$$

Variable $x_{EF\ ij}^o$ expresses whether demand *o* uses lightpath *(E, F)* on edge *(i, j)* or not.

$$y_{ij}^{EF} \in \{0,\ 1\},\ \forall (i,j) \in A,\ (E,F) \in L \tag{5.4}$$

Variable y_{ij}^{EF} indicate whether lightpath *(E, F)* uses edge *(i, j)* or not.

$$y_{ij} \in \{0,\ 1\},\ \forall (i,j) \in A \tag{5.5}$$

Variable y_{ij} expresses whether edge *(i, j)* is used by the routing or not.

We use the same constants and calculated constants defined in 4.1.

5.2 Objective function

Minimize:

$$\alpha \cdot \sum_{\forall (i,j) \in A} y_{ij} + (1-\alpha) \cdot \sum_{\forall (E,F) \in L} p_{EF} \tag{5.6}$$

The objective function expresses that the routing cost (including network resources) should be minimized together with the total of signal powers. If we want to minimize

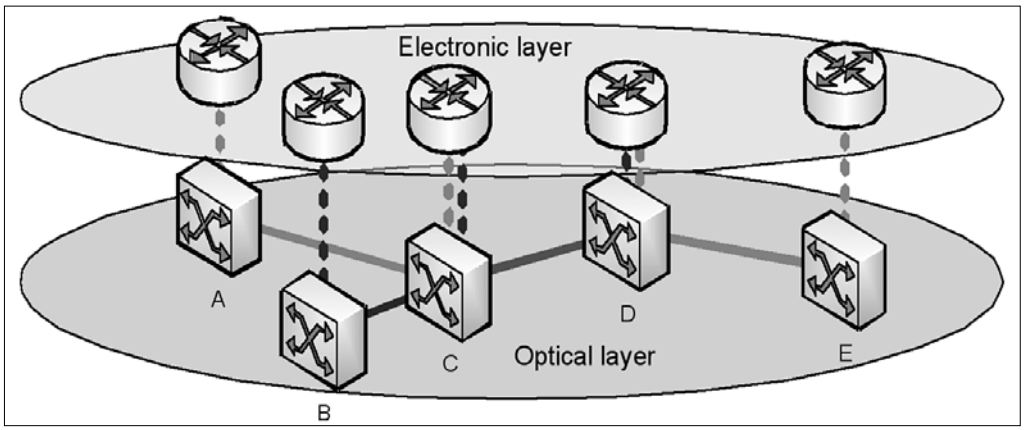


Figure 4.
There are two demands (A-E and B-D). Altogether 4 lightpaths (A-C, B-C, C-D and D-E) are allocated. Grooming is applied on lightpath C-D.

the sum cost of the routing, constant cost factors should be assigned to each edge. Constant α decides whether optimization emphasis is on minimal routing cost (α is close to 1) or on minimal signal power (α is close to zero).

5.3 Constraints

$$\sum_{\forall(i,j)\in pl,\ (E,F)\in L} p_{ij}^{EF} \le 1,\ \forall pl \in PhyLinks \tag{5.7}$$

$$x_{EF\ ij}^{o} \le y_{ij}^{EF} \le y_{ij},\ \forall o \in O,\ i,j \in V,\ (E,F) \in L \tag{5.8}$$

$$y_{ij}^{EF} \le \sum_{\forall o\in O,} x_{EF\ ij}^{o},\ \forall (i,j) \in A,\ (E,F) \in L \tag{5.9}$$

$$y_{ij} \le \sum_{\forall (E,F)\in L} y_{ij}^{EF},\ \forall (i,j) \in A \tag{5.10}$$

$$p_{ij}^{EF} \le y_{ij}^{EF},\ \forall i,j \in V,\ (E,F) \in L \tag{5.11}$$

$$\sum_{\forall j\in V^{-i}} p_{ji}^{EF} - \sum_{\forall k\in V^{i\to}} p_{ik}^{EF} = \begin{cases} -p_{EF} & \text{if } i = E \\ 0 & \text{if } i \notin \{E,F\}, \\ & \forall i \in V,\ (E,F) \in L \\ +p_{EF} & \text{if } i = F \end{cases} \tag{5.12}$$

$$\sum_{\forall j\in V^{-i}} \sum_{\forall (E,F)\in L} x_{EF\ ji}^{o} - \sum_{\forall k\in V^{i\to}} \sum_{\forall (E,F)\in L} x_{EF\ ik}^{o} = \begin{cases} -1 & \text{if } i = s^{o} \\ 0 & \text{if } i \notin \{s^{o},t^{o}\},\ \forall i \in V,\ o \in O \\ +1 & \text{if } i = t^{o} \end{cases} \tag{5.13}$$

$$\sum_{\forall E,\ F} y_{ij}^{EF} \le 1,\ \forall (i,j) \in A \tag{5.14}$$

$$\sum_{\forall o\in O,} \sum_{(E,F)\in L} x_{EF\ ij}^{o} \cdot b^{o} \le B,\ (i,j) \in A \tag{5.15}$$

$$\sum_{\forall (i,j)\in A_{pw},} y_{ij}^{EF} \cdot len_{PhyNode} + \sum_{\forall (i,j)\in A_{pl}} y_{ij}^{EF} \cdot len_{ij} \le \le L(p_{EF}) = L_{c} \cdot p_{EF} \cdot P_{pl\ lin}^{max},\ (E,F) \in L \tag{5.16}$$

5.4 Explanation

Constraint (5.7) explains that the total power of lightpaths traversing a physical link or fiber (denoted by *pl*) cannot exceed the maximum allowed power of that link. In constraint (5.7) we calculate sum of the power of lightpaths going through those edges that belong to physical link *pl*. Constraint (5.8) is straightforward: it expresses that edge *(i,j)* is used by lightpath *(E,F)* if any of the demands – multiplexed into lightpath *(E,F)* – uses that edge. Similarly it also tells that edge *(i,j)* is used by the routing if any of the lightpaths use that edge. Constraint (5.9) states that edge *(i,j)* is used by lightpath *(E,F)* only if it is used by at least one demand. I.e., lightpath *(E,F)* does not use unnecessarily edge *(i,j)*. Similarly constraint (5.10) expresses that edge *(i,j)* is used by the routing only if it is used by at least one lightpath, i.e., a lightpath is not created unnecessarily. Constraints (5.9) and (5.10) are optional, since these rules are implicitly expressed by the objective function. Constraint (5.11) simply means that if the power of a lightpath on an edge is greater than zero, then that edge is used by the lightpath. Constraint (5.12) assures that the signal power of a lightpath is the same along the whole path. Constraint (5.13) expresses flow conservation constraint for demands. Constraint (5.14) assures that each edge is used by at most one lightpath. Constraint (5.15) expresses the grooming constraint, i.e., the sum bandwidth of multiplexed demands cannot exceed wavelength capacity. Constraint (5.16) expresses the relation between the physical distance traversed by the lightpath and the signal power of the lightpath.

6. Benefits of the algorithm

It is a very hard task to illustrate the efficiency of the algorithm since it gives obviously better results than the traditional RWA algorithms. This is due to the additional degree of freedom, namely, the tunability of the signal power. In this section we illustrate some of the benefits of the algorithm having in mind that for different input parameters the results would be slightly different.

In our simulations we used the well-known COST 266 reference network *(Fig. 5)*. Since this network is a long haul network we have decreased the lengths of the links by 25% to get a metro size optical network. The nodes are fully optical nodes and the signal can not be 3R regenerated or converted into the electronics once it is in the optical layer. To demonstrate the advantage of the

Figure 5.
COST 266 European reference network topology

proposed method we have introduced the concept of "maximum routed demands". This means that we have randomly generated a certain number of demands, a traffic matrix. If these demands could be routed, we increase the number of demands, e.g. the size of the traffic matrix, and route it again. This process continues until it is not possible to route more demands anymore. This way it is possible to find the maximum number of demands which can be routed. The bandwidths of the demands were equal with the capacity of one channel. The source and destination pairs were chosen randomly. We used the single layer routing scheme. The constants of the routing algorithm were as described in Section 4.

The absence of solution can have two reasons: the RWA does not succeed or the distance between the source and destination node is too long i.e. the signal quality will be inadequate. It has to be mentioned that the proposed algorithm finds the global optimum of the routing problem which is an NP-hard problem. Therefore in some cases to find the maximum demands which can be routed takes long time, approximately one week for the COST 266 network with 8 wavelengths ($W=8$) and $n=8$, where n means the maximum allowed deviation of signal power from the traditional power allocation scheme (see Equation (4.7)).

The "maximum routed demands" means the number of successfully routed demands from a randomly generated demand set. If a certain number of demands could be routed, we increase the number of demands and route it again. This process continues until it is not possible to route more demands anymore. This way it is possible to find the maximum number of demands which can be routed. The bandwidths of the demands were equal to the capacity of one channel. The source and destination pairs were chosen randomly. This timescale problem is not a significant drawback of the proposed algorithm since in real networks this kind of routing problem will not occur. Finding the global optimum (e.g., for COST 266 network with 8 wavelengths, $n=1.5$ and 60 demands), takes approximately 10 minutes, which is a really fast RWA solution.

We compared the proposed algorithm with the traditional RWA algorithm *(Fig. 6 and 7)*. On the y-axis the maximum number of routed demands is depicted, while on the x-axis the used routing schemes. *RWA* means that we used the traditional routing scheme where each channel has the same signal power. The *n=1* routing scheme is similar to the *RWA* routing scheme. The only difference is that in case of *n=1* the channel powers can be lower than the average of the powers. In *RWA* case this variation is not allowed. In *n>1* cases we used the proposed routing algorithm with *n* equal to the depicted numbers.

The result marked as "MAX" is the number of maximum routed demands in case when physical effects are neglected. The *scale parameters* mean that we changed the lengths of the used network link by multiplying the original lengths with the scale parameter. In Fig. 6 the scale is 1, i.e., we used the original link lengths (geographical distances). In Fig. 7 the scale parameter is 1.25.

In Fig. 6-7 it can be seen that the traditional RWA algorithm can route 19 and 1 demands, respectively. While by increasing the n-factor more and more demands can be route until we reach a limit, where the RWA problem is infeasible in itself (without considering physical effects).

The results lead us to a conclusion that just a small amount of n-factor increases highly increase the number of maximally routed demands. In *Table 1* we have depicted the corresponding channel powers for different n-factor values in mW and dBm. As it is to be seen these values are much lower than the Brilluoin-threshold.

To investigate the dependency of the proposed method on the number of wavelengths we made simulations using the COST266 network topology and different wavelength numbers (see *Fig. 8*). The figure shows that while increasing the number of channels the maximum number of routed demands is increasing. This behavior is as expected when solving the RWA problem. The interesting property is that if we double the number of wavelengths and the n-factor is high enough, the maximum number of routed demands is more than double in each case. This behavior is due to the way how the proposed algorithm works.

If we have more wavelengths, there are more possible variations how the signal power can be allocated. Consequently, if the number of wavelengths is increased, the performance of the proposed algorithm will improve. However, as it was mentioned before, for higher number of wavelengths (32-64) to find the maximum number of demands which can be routed takes very long time, since it needs many tests to find the exact number of demands which can be routed. The other timescale problem occurs when the number of demands is very close to the maximum number of demands which can be route. This kind of simulations can have long running time, more than

	fix	1	1,2	1,4	1,6	1,8	Brillouin threshold
Pimax (mW)	1,25	1,25	1,5	1,75	2	2,25	~ 5
Pimax (dBm)	0,96	0,96	1,76	2,43	3,01	3,52	~ 7

Table 1.
n-factor values in mW and in dBm respectively

a week. In other cases the running time of the algorithm has a timescale of minutes even for higher number of wavelengths.

7. Conclusion

In this paper we presented new RWA algorithms where the power of WDM channels can be adjusted. Our proposed algorithms perform joint optimization of routing (RWA) and of determining signal powers of WDM channels. The proposed methods can be used in existing WDM optical networks where the nodes support signal power tuning.

Figure 6.
Maximum number of routed demands vs. n-factor parameter in case of COST 266 topology

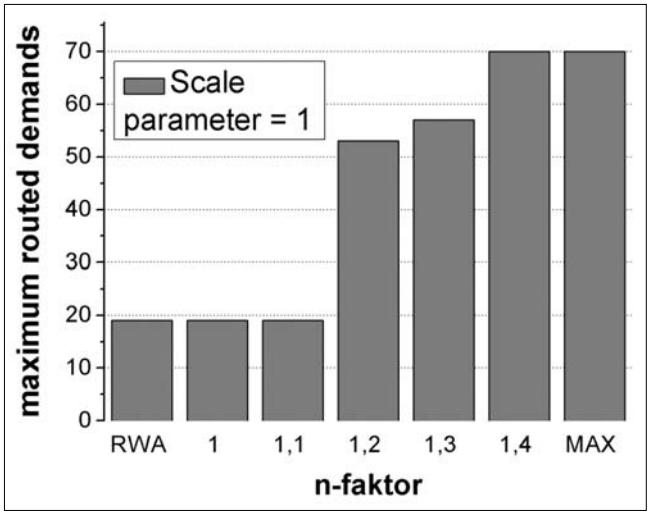
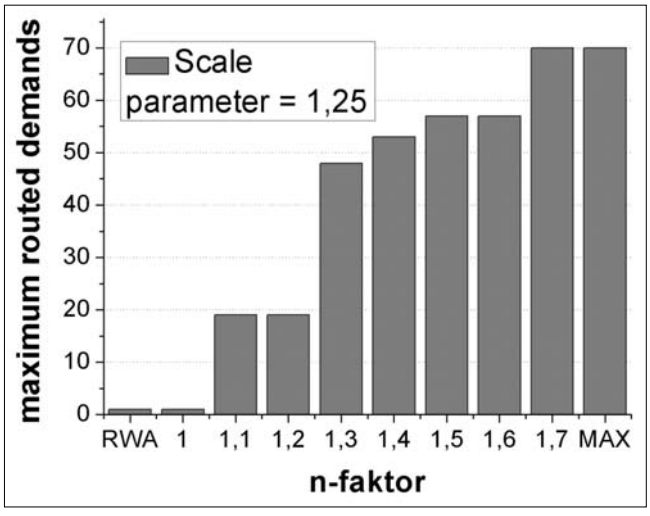


Figure 7.
Maximum number of routed demands vs. n-factor parameter in case of COST 266 topology, scale 1.25



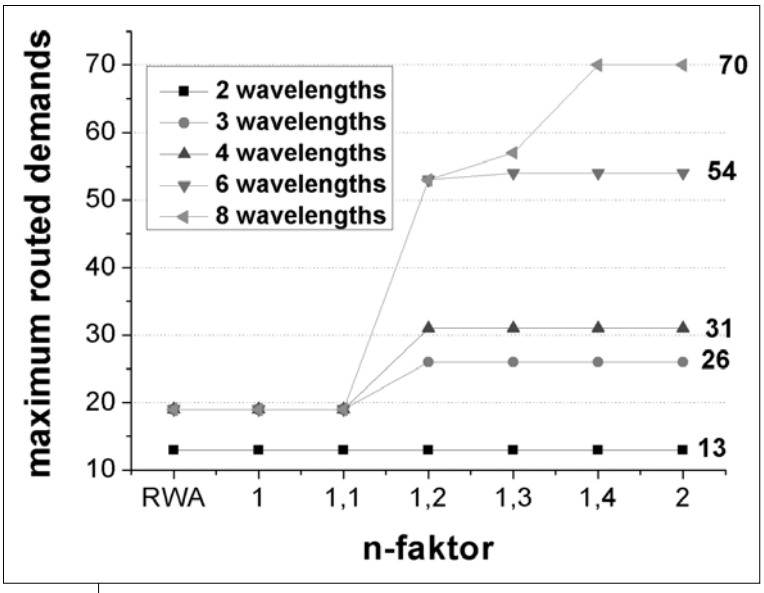


Figure 8.
Maximum number of routed demands versus n-factor parameter in case of COST 266 topology, for different wavelength numbers

We gave the exact ILP formulation of the problems to find the global optimum. In the simpler single layer case it is not allowed to use the electronic layer at all along a path, except for the source and destination nodes, while in the more complex multilayer case electronic layer can perform 3R regeneration, grooming and wavelength conversion.

In the second case we carry out full joint optimization of RWA with grooming (according to a given demand set) and with determining the power of lightpaths (observing physical constraints). The multilayer optimization proved to be too complex for even small networks, while the single layer ILP formulation is practically applicable for moderate size networks.

However it is still worth looking for fast heuristic approaches. For such heuristic methods our ILP based optimal solutions can be regarded as a baseline for comparison.

Authors

SZILÁRD ZSIGMOND has received M.Sc. (2004) degree in physics from the Budapest University of Technology and Economics (BME), Hungary, where he is currently working toward his Ph.D. degree at the Department of Telecommunications and Media Informatics (TMIT). His research interests focus on all optical networks, physical effects of the optical networks, PMD, optical switching architectures, impairments constraint based routing. He is author of several refereed scientific publications. He has been involved in numerous related European and Hungarian projects including 291, NoE e-Photon/ONe and NoE e-Photon/ONe+; CELTIC PROMISE; NKFP.

MARCELL PERÉNYI received his M.Sc. degree in Computer Science from the Budapest University of Technology and Economics, Hungary, in 2005. He is currently a Ph.D. student at the same university and a member of the High Speed Network Laboratory. He has participated in several research project

supported by the EU. His research interests include simulation, optimization and planning of optical networks, as well as traffic identification and analysis in IP networks.

TIBOR CINKLER has received M.Sc. (1994) and Ph.D. (1999) degrees from the Budapest University of Technology and Economics (BME), Hungary, where he is currently associate professor at the Department of Telecommunications and Media Informatics. His research interests focus on optimisation of routing, traffic engineering, design, configuration, dimensioning and resilience of IP, Ethernet, MPLS, ngSDH, OTN and particularly of heterogeneous GMPLS-controlled WDM-based multilayer networks. He is author of over 130 refereed scientific publications and of 3 patents. He has been involved in numerous related European and Hungarian projects including ACTS METON and DEMON; COST 266, 291, 293; IP NOBEL I and II and MUSE; NoE e-Photon/ONe and NoE e-Photon/ONe+; CELTIC PROMISE; NKFP, GVOP, ETIK, etc; and he is member of ONDM, DRCN, BroadNets, AccessNets, IEEE ICC and Globecom, EUNICE, CHINACOM, Networks, WynSys, ICTON, etc. Scientific and Programm Committees. He has been guest editor of a Feature Topic of the IEEE ComMag and reviewer for many journals.

References

[1] XC-VXL-10G/2.5G Cross Connect Cards for the Cisco ONS 15454 SDH MSPP, Cisco ONS 15454 60G/5G High-Order/Low-Order XC-VXC Cross-Connect Card http://www.cisco.com/

[2] LambdaXtremeTM , Lambda Unite MMS datasheet http://www.alcatel-lucent.com/wps/portal

[3] MARCONI MHL 3000 CORE datasheet http://www.ericsson.com

[4] OptiX BWS 1600G DWDM datasheet http://www.huawei.com/

[5] LambdaDriver WDM Tunable Cards datasheet http://www.huawei.com/

[6] LambdaDriver Tunable 10GE WDM cards, TM-DXFP20T and TM-DXFP35T DM Tunable Cards http://www.mrv.com

[7] N. Wauters, P. Demister, "Design of the Optical Path Layer in Multiwavelength Cross-Connected Networks", IEEE Journal on Selected Areas in Communications, Vol. 14, No.5, June 1996. pp.881–892.

[8] R. Ramaswami, K.N. Sivarajan, "Routing and Wavelength Assignment in All-Optical Networks", IEEE Transaction on Networking, Vol. 3, No.5, 1995. pp.489–500.

[9] http://vpisystems.com/

[10] B. Ramamurthy, D. Datta, H. Feng, J.P. Heritage, B. Mukherjee, "Impact of Transmission Impairments on the Teletraffic Performance of Wavelength-Routed Optical Networks," IEEE/OSA J. Lightwave Tech., Vol. 17, No.10, 1999. pp.1713–1723.

[11] Tony Antony, Ashwin Gumaste, "WDM Network Design", Cisco Press, Februar 7, 2003.

[12] T. Cinkler et al., "Configuration and Re-Configuration of WDM networks" European Conference on Networks and Optical Communications, NOC'98, Manchester, UK, 1998.

Determining the optimal signal power based on physical effects in CWDM optical networks

ÁRON SZABÓ, SZILÁRD ZSIGMOND

Department of Telecommunication and Media Informatics, Budapest University of Technology and Economics

{aron.szabo, zsigmond}@tmit.bme.hu

Keywords: CWDM, DWDM, Q-factor, SRS, GVD, RIN

To increase the signal power in optical networks is crucial to extend the maximum transmission distance. However, the nonlinear behavior of the optical fiber limits the signal power e.g. the span of the all-optical network. This paper presents an analytical model and calculation results for the signal quality degradation in an 8-channel and in an 18-channel, 2.5 Gbps coarse wavelength-division-multiplexing (CWDM) system. Based on the proposed model and performed analysis we give the optimal value of the signal power at the transmitter point. The modeling of chromatic dispersion and Raman-scattering, the two main constraints of CWDM optical networks is also presented in detail.

1. Introduction

Today it is becoming increasingly important to transmit data as close to the users as possible, but at a lower price, offering a convenient bit rate. It is mostly the tendency in metropolitan area networks (MANs). Coarse wavelength division multiplexing (CWDM) standard is suitable for this purpose, using the 1270 nm-1610 nm band with 20 nm channel spacing for 18 channels. In most cases only the top wavelength band is built in CWDM systems, using 8 channels from 1470 nm to 1610 nm because of the fiber's OH- attenuation peak at 1383 nm and the lack of high-quality lasers near 1400 nm.

Nonlinear optical effects occurring in wavelength-division-multiplexed systems have been studied extensively by previous literature, but so far, to the best of our knowledge, there has been no study that summarizes the impacts of the significant linear and nonlinear optical effects in CWDM systems and derives the optimal inserted signal power in view of them.

The paper presents an analytical model and calculation results for the signal quality degradation in 8-channel and 18-channel, 2.5 Gbps point-to-point CWDM links due to physical effects. We performed the calculations for 3 different fiber lengths: 60 km, 100 km, 140 km. We used the results to determine the highest input signal powers at which the signal quality remains convenient. We regarded these powers as optimal inserted powers.

2. The theoretical model

Nonlinear optical effects occurring in the DWDM systems have been investigated in [1,2]. In the case of CWDM systems only the stimulated Raman-scattering (SRS) and the stimulated Brillouin-scattering (SBS) have significant impacts because of the shorter fiber lengths and the lower signal powers [2,3]. Since most CWDM systems work at 2.5 Gbps per channel bit rate, this bit rate is assumed in the model. The polarization mode dispersion (PMD) and the polarization-dependent loss (PDL) are neglected because of the relatively low bit rate [4]. Dispersion compensation units (DCUs) are widely used in DWDM systems, but ignored in CWDM systems, consequently group velocity dispersion (GVD) is important and affects differently each channel. The studied architecture is shown in *Fig. 1*. The main purpose of our optimization is to reach the highest power at the transmitter (point P) at which the signal quality remains convenient at the receiver (point Q), see Figure 1.

The bit 1 and the bit 0 signal levels are assumed to be Gaussian random variables with μ_1 and μ_0 mean values and their standard deviations are σ_1 and σ_0. The signal quality at the receiver point is described by the Q-factor:

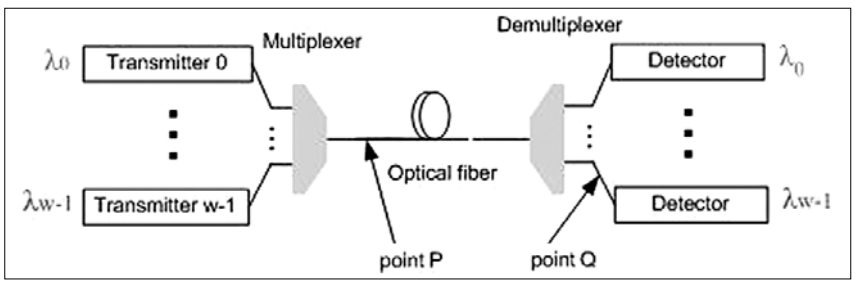
$$Q = \frac{\mu_1 - \mu_0}{\sigma_1 + \sigma_0}, \qquad (1)$$

For the calculations we used the numerical data of the prevalent ITU-T G. 652 single mode optical fiber.

2.1 The impact of the transmitter intensity noise

Beyond physical effects of propagation the model considers the transmitter intensity noise. Amplifiers are

Figure 1. The studied architecture



not widespreadly used in CWDM systems so amplifier spontaneous emission (ASE) is not involved in the model [5]. But the noise produced by the relative intensity noise (RIN) of the laser must be considered. The RIN is power-dependent and we used the -120 dB/Hz maximum value (RIN_{dB}) at 0 dBm signal power, which is prevalent for CWDM transmitters [6]. A method shown in [7] is used to calculate the dimensionless noise ratio (σ) of a signal:

$$\sigma = \sqrt{B_c \cdot 10^{RIN_{dB}/10}}, \quad (2)$$

where B_c is the receiver electrical bandwidth. Denoting the signal powers 1 and 0 by P_1 and P_0, taking the prevalent 7.4 dB extinction ratio for P_1/P_0 at the transmitter point and assuming that the power dependence of the RIN [1/Hz] is proportional to $1/P^3$ [8], the RIN-related noise of the signal levels can be readily calculated.

2.2 The propagation-related effects

The propagation-related effects include the chromatic dispersion and the stimulated Raman-scattering. The stimulated Brillouin-scattering limits the power so that the total inserted power is backscattered above the Brillouin-threshold. Since many technologies exist to increase the Brillouin-threshold thus present calculation is executed both, below and over the calculated Brillouin-thresholds.

2.2.1 The effect of the GVD

To calculate the effects of GVD onto the signal we sampled it at the half of the bit period at the ending point of the fiber. The chromatic dispersion changes the original signal shape, therefore the sampled signal level differs from the maximum of the level inserted into the fiber at the starting point.

The change of the signal shape is calculated using the equations shown in [1]:

$$U(z,T) = \quad (3)$$

$$(1/2\pi)\int_{-\infty}^{\infty} U(0,\omega)\exp\left[(i/2)\beta_2(\omega)\omega^2 z - i\omega T\right]d\omega,$$

where $U(z,T) = \left(1/\sqrt{P}\right)\exp(\alpha z/2)A(z,T)$ (4)

is the attenuation-normalized envelope, z is the distance from the starting point of the fiber,

$$T = t - z/v_g \quad (5)$$

is the time measured at the coordinate system moving with the envelope, v_g is the group velocity, P is the power inserted into the investigated channel, α is the attenuation,

$$U(0,\omega) = \int U(0,T)\exp(i\omega T)dT \quad (6)$$

is the Fourier-transform of the inserted signal at $z=0$. $\beta_2(\omega)$, which is the second term in the Taylor series of the mode propagation constant, is calculated using the disclosed $D(\lambda)$ dispersion parameter of the ITU-T G.652 fiber [9]:

$$\beta_2 = -\frac{\lambda^2}{2\pi c}D(\lambda) \quad (7)$$

For further calculations the knowledge of the $U(z,T)$ signal shape is required. For directly modulated semiconductor lasers the signal shape is well approximated by the super-Gaussian function [1]:

$$U(0,T) = \exp\left[-\frac{1+iC}{2}\left(\frac{T}{T_0}\right)^{2m}\right], \quad (8)$$

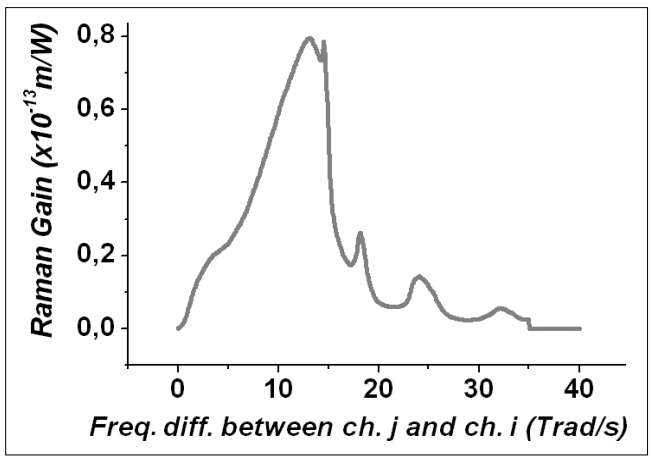
where T_0 is the half-width at the $1/e$ intensity point, C is the chirp parameter characterizing the time dependence of the spectral power density of the laser and the m parameter controls the degree of the pulse edge sharpness. Using disclosed transmitter data [10], a realistic estimation is $m=3$. In our calculation we use a typical value of C=-3,6 [11]. We suppose that the chirp parameter is independent of the modulation frequency and the intensity [12].

Because of the non-return-to-zero (NRZ) modulation, the shape of a given bit pulse is modified not only because of its dispersion but the dispersion of the previous and the subsequent bit. We assumed that the impacts of the pulses further from the investigated pulse than 2 bit times are neglected. For both bit 0 and bit 1 signals we performed the calculation in the cases of each possible adjacent bit, assuming these bit sequences to have the same probability. In the case of bit 1, the 010, 011, 110, 111 sequences are investigated and they have one by one 1/4 probabilities, similarly in the case of bit 0, the 000, 001, 100, 101 sequences are investigated and have also one by one 1/4 probabilities. Simulating NRZ modulation, the group velocity dispersion related μ_{1GVD} and μ_{0GVD} mean values are calculated at the ending point of the fiber. In our model we use the previously released approximation [13] that the chromatic dispersion affects the mean values but leaves the standard deviations invariably, see Equation (1).

2.2.2 The stimulated Raman-scattering (SRS)

Stimulated Raman-scattering is an inelastic scattering process in which higher energy photons of the pump wave scatter on the medium molecules producing lower energy photons and optical phonons. The original beam is called pump wave, the beam containing the lower energy photons is called Stokes wave. If there are photons at the lower energy state, SRS becomes an amplification process, [1]. In the case of CWDM systems it happens if the frequency difference of two channels is in the Raman gain bandwidth *(Fig. 2)*. This way energy is transmitted from the higher energy channels into the

Figure 2. The Raman gain spectrum



lower energy ones. Amplification occurs in the parts of the fiber where the bit 1 signal of the pump wave overlaps the bit 1 signal of the Stokes wave. In the case of SRS the Stokes waves have the same direction as the pump waves therefore noise is produced both in the Stokes waves and in the pump waves, deteriorating the signal quality.

Investigating the interaction of only 2 channels with narrow channel spacing, the interaction of the pump and the Stokes wave has already been studied in [14]. This model is ideal for DWDM-related calculations, where the channel spacing is approximately 1 nm. The model must be modified in the case of CWDM systems because of the 20 nm channel spacing.

The integral formula which describes the SRS crosstalk includes a $g_R'\Delta f_{ji}$ factor where i denotes the number of a pump channel and j denotes the number of a Stokes channel, Δf_{ji} is the frequency difference between the pump and the Stokes wave, g_R' is the Raman-gain slope, which is the derivative of the Raman-gain spectrum near $\Delta\omega_{ji}=0$.

In the case of CWDM systems, the $g_R'\Delta f_{ji}$ factor must be replaced by $g_R(j,i)$, which is the Raman-gain value at $\Delta\omega_{ji}$. Figure 2 shows the Raman-gain spectrum of the ITU-T G. 652 single mode fiber using 1550 nm pump wavelength. The gain value is inversely proportional to the pump wavelength [1].

In [14], the depletion of the pump channel caused by the SRS is characterized by a Gaussian random variable $x(z,t)$, where z and t denote the distance from the starting point of the fiber and the time. Approximating the waveforms by NRZ-modulated rectangular shapes, the deviation and the mean value of $x(z,t)$ denoted by σ_x and μ_x can be calculated [14]. In this case, z equals to the fiber length.

In the current many-channel case we used the approximation that the interaction of the channels can be discussed by each possible combination of pairs of pump waves and Stokes-waves. In this case, σ_x and μ_x are replaced by σ_{xji} and μ_{xji} representing the interaction of channel j and channel i. For a given channel i we summarized the impact of all the other channels to determine the evolution of the power in channel i:

$$\mu_{xi}=\sum_{\substack{j=1\\ j\neq i}}^{W}\mu_{xji}-\sum_{\substack{j=1\\ j\neq i}}^{W}\mu_{xij} \tag{9}$$

$$\sigma_{xi}^2=\sum_{\substack{j=1\\ j\neq i}}^{W}\sigma_{xji}^2+\sum_{\substack{j=1\\ j\neq i}}^{W}\sigma_{xij}^2 \tag{10}$$

where W is the number of the channels. In (9) the mean values of channels that describe the power scattered from channel i are positive, and the mean values that describe the power scattered into channel i are negative. We assume that each impact can be described by independent probability variables; Equation (10).

Using (9) and (10), the real effect of SRS on the signal level and noise, μ_{1SRS} and σ_{1SRS} respectively, can be calculated for each channel [14].

2.3 Calculating the Q-factor

Q-factors are calculated one by one for each channel to describe the total impact of the GVD, the RIN and the SRS at given fiber lengths as functions of the inserted power. They are calculated using the approximation that the GVD and the SRS act independently. In our model, the same signal power was inserted into each channel. The attenuation is taken into account by a multiplication with $e^{-\alpha L}$, where α and L denote the attenuation and the fiber length.

As it can be seen by the next equations, in this model the attenuation has no effect on the Q-factor:

$$Q_{tot}=\frac{\mu_{1SRS}\mu_{1GVD}e^{-\alpha L}-\mu_{0GVD}e^{-\alpha L}}{\sqrt{\sigma_{1RIN}^2+\sigma_{1SRS}^2}e^{-\alpha L}+\sigma_{0RIN}e^{-\alpha L}}=$$
$$=\frac{\mu_{1SRS}\mu_{1GVD}-\mu_{0GVD}}{\sqrt{\sigma_{1RIN}^2+\sigma_{1SRS}^2}+\sigma_{0RIN}} \tag{11}$$

We note that irrespectively of this the receiver sensitivity limits the minimum received power.

Furthermore we define the Q_{SRS} to describe the effect of SRS alone on the signal quality:

$$Q_{SRS}=\frac{\mu_{1SRS}}{\sigma_{1SRS}} \tag{12}$$

and similarly

$$Q_{GVD,RIN}=\frac{\mu_{1GVD}-\mu_{0GVD}}{\sigma_{1RIN}+\sigma_{0RIN}} \tag{13}$$

In (12) we used that the SRS has effect only on the signal level 1, and in (13) it is used that the GVD have effect only on the mean values, while SRS modifies only the deviations.

3. Calculation Results

In this section we present the calculation results for the Q factors of the 8-channel CWDM systems that uses the upper CWDM wavelength band and also for 18-channel systems that uses the whole band. We performed the calculations for 3 different fiber lengths and present the Q factor values of each channel at the receiver point as a function of the inserted power.

We assume that the signal quality remains convenient if $Q\geq 14$ for each channel. This choice is made with the estimation that the signal quality is convenient if $Q\geq 7$, and if the Q factor is reduced by the physical effects to $Q\geq 14$, other effects occurring in the network cannot reduce it below 7. Using this assumption, the maximum inserted power for a given system configuration is defined as the highest inserted power at which $Q\geq 14$ for each channel. The channel for which $Q=14$ at this power is called the most power-sensitive channel. *Table 1* summarizes the main parameters of the optical system used for the calculations.

Fig. 3 shows the general behavior of $Q_{GVD,RIN}$, Q_{SRS} and Q_{tot} as functions of the inserted power. The example shows the Q factor values of the 1470 nm channel in the 8-channel system with 140 km fiber length.

Bit rate	2,5Gbps
Wavelength band of the 8-channel CWDM system	$1470nm-1610nm$
Wavelength band of the 18-channel CWDM system	$1270nm-1610nm$
Channel spacing	$20nm$
Speed of light in vacuum	299792458 m/s
Effective core area	$80\mu m^2$
Maximum RIN at 0 dBm	$-120dB/Hz$
Optical bandwidth of the receiver	$10GHz$
Electrical bandwidth of the receiver (B_e)	$1,75GHz$
Fiber length (L)	$60km$, $100km$, $140km$
Chirp parameter of the transmitter (C)	-3,6
Super-Gaussian edge sharpness parameter (m)	3
Extinction ratio (P_1/P_0)	7,4 dB
Dispersion parameter $D(\lambda)$	Fiber specification [9]
Raman gain (g_R)	Fig. 2.

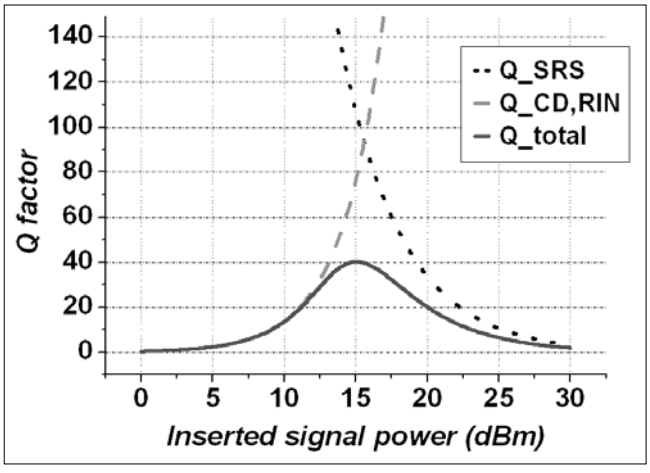
Table 1. The main parameters of the optical system

As it can be clearly seen, increasing the inserted power produces increasing $Q_{GVD,RIN}$, because the proportion of the RIN to $1/P^3$ produces decreasing $\sigma_{1\,RIN}$ and $\sigma_{0\,RIN}$, while the effect of the GVD does not depend on the inserted power but it is a function of the fiber length and the wavelength of the channel. However, if the inserted power is high enough, the effect of the SRS becomes significant and deteriorates the signal quality, therefore Q_{SRS} decreases. Taken all round, a maximum value of the Q factor at a given power exists for each channel in each system configuration at different power values.

Fig. 4 shows the power-dependence of the Q factors calculated for each channel in 8-channel and 18-channel systems with 60 km, 100 km and 140 km fiber length. We show the Q factor values of 3 channels in the case of the 8-channel systems and the Q factor values of 4 channels in the case of the 18-channel systems including the most power-sensitive channel in each case.

In the case of the 8-channel systems the most power-sensitive channel has a relatively low wavelength. The inserted powers that belong to the maximum values of the Q factors for different channels are changing even for the same fiber length, but they are all near 15 dBm. This aberration is caused by the wavelength dependence of the GVD and the SRS.

Figure 3.
Q factor dependency from chromatic dispersion, relative intensity noise and Raman scattering



In the 18-channel case the wavelength of the most power-sensitive channel is in the medium wavelength domain in contrast to the 8-channel case. The power values that belong to the maximum Q factors of different channels differ for the same fiber length in this case too, but the power values are on average higher than in the 8-channel case, they are near 20 dBm.

We defined the optimal signal power to be the highest power at which $Q \geq 14$ for each channel. *Table 2* shows these power values both for the 8-channel and the 18-channel systems.

	60 km	100 km	140 km
8 channels	21.2 dBm	20.0 dBm	21.0 dBm
18 channels	25.3 dBm	25.1 dBm	26.3 dBm

Table 2. The optimal signal power values for the 8-channel system with different fiber lenghts (Q≥14)

The optimal power values of the 18-channel system are appreciably higher than those of the 8-channel system, which is a direct consequence of the higher channel number. The optimal power values do not increase linearly with the channel number because of the increasing intra-channel interaction.

Finally, we can answer a question: “Can a signal be better after 140 km than after 100 km?” Yes, it can, as it is clearly seen by Figure 4. and by Table 2. The Q factor values and the optimal signal powers are higher than after 100 km in both the 8- and the 18-channel case.

The key is the dispersion of the super-Gaussian signal shape. This shape not only broadens but produces peaks and valleys because of the GVD [15]. For appropriate fiber lengths, e.g. 140 km, a dispersion-induced peak arrives at the ending point of the fiber, enhancing the Q factor considerably.

4. Conclusions

In this paper we presented the dependence of the physical effects and the signal quality on the signal power and the fiber length for 2.5 Gbps CWDM systems. We show by analytical calculations what the optimal signal powers are for different network scenarios. These results are useful tools for network designers for improving their optical network or even redesigning their power budget calculations.

Authors

ÁRON SZABÓ is currently working toward his M.Sc. degree in physics at the Budapest University of Technology and Economics (BME), Hungary. His research is focused on the calculation of physical effects appearing in different optical network architectures, and the computer simulation of fiber lasers. He is doing research in cooperation with the Department of Telecommunications and Media Informatics (TMIT).

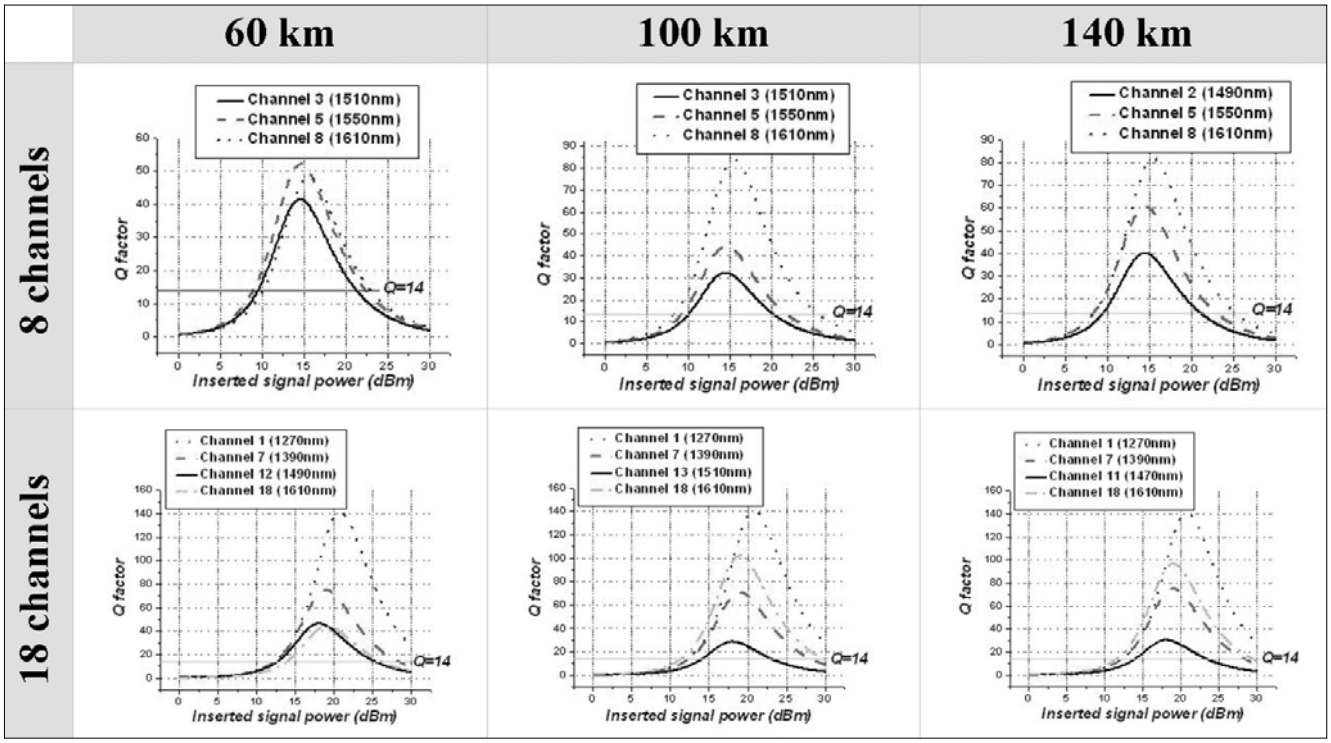


Figure 4. Input power dependence of the Q factor in an 8-channel and in an 18-channel system with 60 km, 100 km, and 140 km fiber length

SZILÁRD ZSIGMOND has received M.Sc.(2004) degree in physics from the Budapest University of Technology and Economics (BME), Hungary, where he is currently working toward his Ph.D. degree at the Department of Telecommunications and Media Informatics (TMIT). His research interests focus on all optical networks, physical effects of the optical networks, PMD, optical switching architectures, impairments constraint based routing. He is author of several refereed scientific publications. He has been involved in numerous related European and Hungarian projects including 291, NoE e-Photon/ONe and NoE e-Photon/ONe+; CELTIC PROMISE; NKFP.

References

[1] Govind P. Agrawal, "Nonlinear Fiber Optics", 3rd edition, Academic Press, 2001.
[2] VPI Transmission Maker User's Manual, VPI Photonics, 1996-2007. http://www.vpiphotonics.com/
[3] Adding/Splitting Nodes With Limited Glass Using CWDM Technologies; Central FL SCTE Chapter Presentation, 1 Nov., 2006.
[4] Vivek Alwayn, "Fiber-Optic Technologies", Pearson Education, Cisco Press, 2004. http://www.ciscopress.com/
[5] W. Shin, I. B. Sohn, B.-A. Yu, Y. L. Lee, S. C. Choi, Y.-C. Noh, J. Lee, D.-K. Ko, "Microstructured Fiber End Surface Grating for Coarse WDM Signal Monitroing", IEEE Photonics Technology Letters, Vol. 19, No.8, 15 April, 2007.
[6] Finisar Product Specification: Multi-rate CWDM Pluggable SFP Transceiver, FWDM-1621-7D-xx; Finisar Corporation, September 2005, Rev. F.
[7] "Impact of transmitter RIN on Optical Link Performance" Maxim High-Frequency/Fiber Communications Group, Application Note, HFAN-9.1.0 Rev 0; 10/04.
[8] L. A. Coldren and S. W. Corzine, "Diode Lasers and Photonic Integrated Circuits", New York, Wiley, 1995.
[9] Drew Perkins, "Dispersion and Skew", IEEE HSSG Interim, 20 September, 2006.
[10] Appointech INC – 2.5 Gbps CWDM laser diode modul, 2005-10. http://www.appointech.com/downloads.html
[11] Aragon Photonics Labs White Paper, (WP001_0100_0307), 2007. http://www.aragonphotonics.com/docs/gc_fichas/doc/56FIPSprxy.pdf
[12] A. Villafranca et al. "Linewidth Enhancement Factor of Semicondutor Lasers: Results from Round-Robin Measurements in COST 288", CLEO07, Baltimore, USA, May 2007.
[13] Bing Xie, Yong Liang Guan, Jian Chen, Chao Lu, "Improvement of dispersion tolerance using wavelength-interleaving and forward error correction", IEEE School, Nanyang Technological University, Singapore; 10 July 2006.
[14] Keang Po Ho, "Statistical Properties of Stimulated Raman Crosstalk in WDM Systems", Journal of Lightwave Technology, Vol. 18, No.7, July 2000, pp.915–921.
[15] Guangqiong Xia, Zhengmao Wu, Jianwei Wu, "Effect of Fiber Chromatic Dispersion on Incident Super-Gaussian Pulse Transmission in Single-Mode Fibers", July 8, 2005., pp.116–120.

Multi Provider Access Networks
Purposes and achievements of MUSE European IST FP6 integrated project (2004-2007)

CSABA LUKOVSZKI

Department of Telecommunication and Media Informatics, Budapest University of Technology and Economics

R+D projects

In serving the growth of the number of subscribers and the magnitude of access services the leading role of the Ethernet-based technologies in the access networks is indisputable nowadays. However, new and unified solutions are required in the range of services, the quality of services and the reliability of networks to keep the costs of investment (CAPEX) and operation (OPEX) low.

In 2003 the European Project MUSE – Multi Service Access Everywhere *(http://www.ist-muse.org/)* – aimed at developing such solutions which are able to serve tens of thousands of users under favorable, low costs. To carry out this task, several European telecommunication institutes (e.g. Alcatel-Lucent, Ericsson, Siemens, Deutsche Telekom, British Telecom, France Telecom, Poland Telecom and other research institutes) joined to perform such research and technical deployments and developments which are able to expand the services of network (such as quality, reliability, security, etc.) by using primarily the already existing elements and to improve the ability of network management by utilizing the already existing tools and solutions. In design of the probable and feasible solutions a great deal of architectural aspects and service provider roles had to be taken into account.

These are depicted in *the figure below*.

The Department of Telecommunications and Media Informatics *(http://www.tmit.bme.hu/)* of Budapest University of Technology and Economics took its share of this work, certainly in the topics closely related to applied and basic research.

The participation of the department focused essentially on the performance analysis and network management of Ethernet networks however, we have proposed research results e.g. of addressing and switching schemes of IPv6.

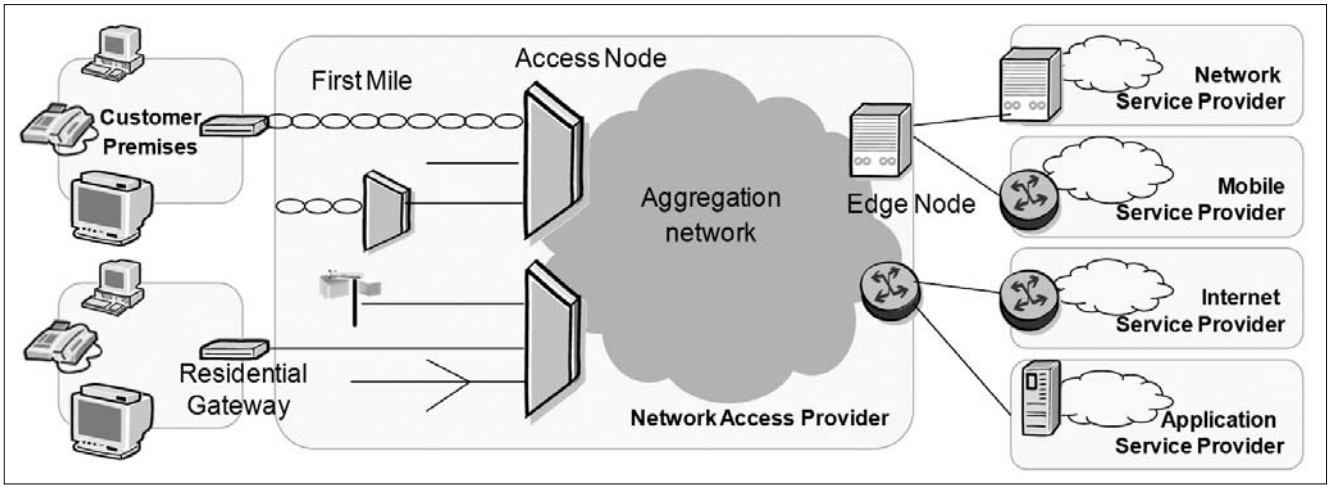
Our main achievements were attained in the traffic optimization of the Ethernet access networks which takes criteria into account such as quality of services and reliability of networks, and in modeling of DSL packet-level scheduling, as well as in the performance analysis of spanning-tree protocol process. From the first mentioned subject an entire thesis has been materialized, but the other subjects are also taken place in other dissertations.

Besides the research subjects, the demonstration of multi layer topology discovery methods of heterogeneous access networks is also important to mention, which traveled through Europe in the last two years (2006-2007) of the project. The demo was shown in open and exclusive presentations in Vienna, Stockholm and Antwerp.

Beyond the research goals – if we would like to evaluate the results of the project – it is also an important aspect that the European participants, including ourselves, were able to receive a good picture about the future goals of their partners and competitors, so the common shaping of the determinant guidelines and policies in telecommunications might be possible in the near future.

To common satisfaction, we took our part from the work in proportion to our role.

MUSE access architecture



Hungarian participation in the Network of Excellence ISIS

MÁRK CSÖRNYEI

Department of Broadband Infocommunication and Electromagnetic Theory, Budapest University of Technology and Economics

R+D projects

The Optical-Microwave Telecommunication Laboratory at the Department of Broadband Infocommunication and Electromagnetic Theory of the Budapest University of Technology and Economics has participated in several successful EU projects during the last couple of years. Recently the laboratory is involved in the IST-ISIS project which is a Network of Excellence supported by the Sixth Framework Programme for Research and Technological Development (FP6) of the European Union.

The project acronym stands for **I**nfra**s**tructures for broadband access in wireless/photonics and **I**ntegration of **S**trength in Europe. ISIS integrates the research activities of 19 organisations from 12 different countries and aims to strengthen European scientific and technological excellence in low cost optical solutions for broadband access, and the merging of wireless and photonic technologies.

The project addresses broadband analogue and digital communication systems like Fibre-to-the-Home (FTTH), the contribution of optical technologies (in the domain of microwave and millimetre-wave photonics) to systems for future fixed and mobile broadband access, low-cost access and edge network equipments, together with advanced wireless sensor network technologies both in microwave and millimetre-wave bands.

Besides scientific work, the cooperation includes significant integrating mechanisms such as organisations of training courses, summer schools, workshops, student and researcher exchanges and development of joint research platforms and tools. On the one hand these events allow faster and more effective penetration of the latest research results into the university education, on the other hand strong, useful contacts and relationships are created between the different European research organisations, which can end up in possible future joint projects and cooperation.

After two years of work one of the most important tasks of the last project period is to ensure the sustainability of research results, aiming at industrial applications in the covered scientific area by opening of photonic-wireless research platforms and test laboratories for European small and medium size enterprises (SME).

More information is available at the project home page: *www.ist-isis.org*.

The Optical-Microwave Telecommunication Laboratory of the Budapest University of Technology and Economics has contributed to the research work of the project by optical network analysis for optimizing subcarrier multiplexed HDTV transmission over the band of 10 Gbps baseband optical transmission and by establishing an IEEE 802.15.4 ZigBee Wireless Sensor Network. Considering the merging of wireless and photonic technologies the main field of research work is now being focused on optical crossconnects between subnetworks of wireless sensor systems. In this case the optical fibre based communication operates as an extension of the wireless network, and thus it enables the remote monitoring of the network or just the substitution of the radio link in case of high wall attenuations in in-door environment.

The research results of the ISIS project have been published in international journal papers and conference proceedings and this publishing activity will hopefully allow fruitful joint research activities in the future, both on national and international level.

Networks 1978–2008: Personal Notes

BY PROF. GYÖRGY LAJTHA

By 1978, the digital transmission was already widespreadly used in telecommunication networks; standards were accepted and maintenance methods were established. After a long period of experimentations, the first stored-program time-division switches were introduced. The attenuation of optical cables gradually decreased and it became already possible to transmit signals of several hundreds of Mbps at less than 5 dB/km. These new technologies required new recommendations and new planning methods.

CCITT, the predecessor of the International Telecommunication Union, recognized this demand and prepared the "General Network Planning Handbook" to facilitate the development and planning. Recommendations related to the applications of the new technologies were prepared by several committees for the 1980 General Assembly at a joint meeting in early fall 1978. Here, one of the leaders of the Paris institute CNET, A. Spizzichino, proposed the idea of organizing a series of symposia in order to permanently follow the new results in research and technology.

He invited a small group of specialists to become members of the organizing committee, and thus the new conference series called Networks was founded. At the first meeting, participants were also Chris Nivert (Sweden), W.E. Falconer (USA), F. Alvarez Casals (Spain), Keith Ward (UK), Anthon Mullen (Ireland), J.C. Lutcheford (Canada), Prof. D. Cagliardi (Italy) and N. Noort (The Netherland). It was decided to hold the conferences tri-annually at different locations. There was also an agreement to keep away business-type presentations and also that the number of participants shall be limited to 400 to make possible an efficient exchange of ideas. The next meeting was scheduled to the following year, with the objective to discuss the program of the first conference to be held in 1980. We started its organization via intense correspondence and a preparation meeting was held where, in addition to the above people also H. Ikeda (Japan) and R. Meisel (Germany) took part.

After the success of the *First International Network Planning Symposium* held in Paris, it was decided to organize the second one in Brighton, England. Its host was Keith Ward, who did an excellent job with the help of the organizing committee, whose members became good friends by that time.

The third conference was held in Florida, USA, in a small holiday village. One of the most important lessons I learned at this conference was that, due to the decreasing prices of the electronics and increasing real estate prices, not only the equipment costs should be optimized but also the installation costs have to be taken into account and the latter can be the key cost component. This trend later continued and had an increasing role in network design.

The 1989 conference was organized by the Spanish team lead by the highly energetic Pietro Caballero who replaced F. Alvarez Casals in the International Scientific Committee. France was represented by M. Peyrade because of Spizzichino's health problems and Germany replaced Meisel by Peter Heuer. This team then continued to work together for a long time and had many interesting technical discussions on topics that otherwise were not publicly discussed.

Then Japan volunteered to host the 1992 conference, also the preparatory meeting was held in Japan which was a good opportunity to get acquainted with some Japanese information technology laboratories and see how much investment was done in research and development, which, according to the hosts, was supposed to have a short return. At this conference we were witnessing the penetration of the computing technologies which influenced the network design, operation and equipment technologies. Here we could first encounter with the virtual reality, with presentation technology that demonstrated the status of the worldwide network and with real-time planning methods that adjusted the traffic routing to the actual traffic and technical conditions.

After a conference which was considered to be highly successful in every respect, our task was to organize the 1994 conference in Budapest. The opening ceremony was held in the historic building of the Pesti Vigadó. With the support of HTE, the Hungarian Scientific Society for Telecommunications, an excursion was organized to Visegrád, a medieval royal town not far from Budapest, and the conference banquet with a folk dance program was held in the building of the National Gallery. There were changes in the organizing committee: W.E. Falconer and Chris Nivert from the original team unfortunately could not be there and Noort was also missing because of his long-term travelling. The USA was therefore represented by Steve Chen and The Netherlands by Mr. Harmsen. The most significant technical event of the conference was the talk of Steve Chen, who showed that tariffs influence the design to the greatest extent due to the continuing decrease of equipment prices. Therefore the capacities should be determined using risk assessment so that all calls are transmitted with a high probability and thus the revenues are maximized.

In 1996 the conference was organized in Sydney, Australia. By that time it became clear that the organization cannot be successful without industrial support. Therefore, an exhibit was organized as a parallel event with 10-12 booths demonstrating telecommunication products. Here we could feel the first time that telecom industry and services are not country-specific and the world is open to international competition, therefore companies showed new product in a very careful way to avoid disclosure of information to their competitors. Thus instead of public announcements some pieces of sensitive information could only be obtained during lunches and coffee breaks.

Next time Italy organized the event in 1998, and the conference was held in the South Italian city of Sorrento. By that time, after 20 years, the international organizing committee was almost renewed, only a few members of the original committee were present, some of them only as presenters or participants. I too transferred my place in the committee to Gyula Sallai in 1994.

The 2002 event was organized in Munich, Germany, by Mr. Gross, the head of application technology at Deutsche Telekom. Here the number of participants hardly reached 300, and it became obvious that the series cannot be continued based on the 1978 principles only. That's why or maybe also for other reasons that there was not enough application to host the meeting overseas, and it was decided to break the tradition of alternating sites in Europe and overseas. The German Electrotechnical Society VDE volunteered to organize the 2004 event in Vienna, Austria. Here the trend of decreasing participation continued but the event nevertheless was useful as well as pleasant.

Unfortunately it became more and more difficult to organize international conferences due to the high competition and there was a feeling that there would be no volunteer for the organization from Europe or North America beyond 2006 as the countries of both continents have already fulfilled their obligations. Therefore, at the organization meeting after the Vienna conference in 2004, it was decided that India would host the next event with the support of the international community. Here the participation was the lowest so far mostly due to the high travel expenses.

Now we arrived to 2008 when a country, Hungary, will host the conference second time.

Editorial Office (Subscription and Advertisements)**:**
Scientific Association for Infocommunications
H-1055 Budapest V., Kossuth Lajos tér 6-8, Room: 422
Mail Address: 1372 Budapest Pf. 451. Hungary
Phone: +36 1 353 1027, Fax: +36 1 353 0451
E-mail: info@hte.hu
Web: www.hte.hu

Articles can be sent also to the following address:
Budapest University of Technology and Economics
Department of Telecommunications
Tel.: +36 1 463 3261, Fax: +36 1 463 3263
E-mail: szabo@hit.bme.hu

Subscription rates for foreign subscribers:
12 issues 150 USD, single copies 15 USD

Publisher: PÉTER NAGY • Manager: ANDRÁS DANKÓ

HU ISSN 0018-2028 • Layout: MATT DTP Bt. • Printed by: Regiszter Kft.