JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

2/2001

Editorial Board

Editorial Advisory Board

ISSN 1509-4553

druk i oprawa: AG Poligrafia, Józefów ul. Matejki 28a
tel. (022) 789 26 08, www.agpoligrafia.com.pl

JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

Preface

In the year 2001 we celebrate the fifty anniversary of the activity of National Institute of Telecommunications. The history of this Institute extends from year 1934, when Professor Janusz Groszkowski founded the first Polish national research institute in telecommunications. Before the second war, the institute had achievements in telephony, radiophony and television. After the second war, in the time of planned, closed economy in Poland, it provided knowledge and expertise in the field of modern telecommunications, preparing Polish equivalents of international solutions of telecommunication devices. In the last decade, National Institute of Telecommunications had to adapt itself to the opening of Polish economy. It started gradually to change its own character, following the liberalisation of the Polish market, looking for the new, long-term tasks.

In the first (invited) paper of this issue written by Andrzej P.Wierzbicki, General Director of the Institute, and entitled *Challenges before National Institute of Telecommunications after fifty years of its activity* the transformation of the Institute is presented. New mission of the Institute, the results of structural changes of the Institute in last five years, an outline of long-term research programme and the challenges before the Institute on the verge of the new millennium are described.

Paweł Szczepański
Editor-in Chief

Invited paper

Challenges before National Institute of Telecommunications after fifty years of its activity

Andrzej P. Wierzbicki

Abstract — In the year 2001 we celebrate the fifty anniversary of the activity of National Institute of Telecommunications under its present name (in Polish: Instytut Łączności). Although we can date earlier the beginnings of the Institute, we will not present here its detailed history that is discussed in other papers. This paper presents rather the transformation of the Institute in last years, resulting from the systemic transformation of Poland. We describe: new mission of the Institute, the results of structural changes of the Institute in last five years, an outline of long-term research programme. Finally, the paper addresses the challenges before the Institute on the verge of the new millennium.

Keywords — National Institute of Telecommunications, mission, research programme.

1. Introduction: new mission of the National Institute of Telecommunications

The history of the National Institute of Telecommunications extends from the year 1934, when professor Janusz Groszkowski, a world-known specialist in radiocommunications, founded the first Polish national research institute in telecommunications. Before the second world war, the institute had several achievements in telephony, radiophony and television. Experimental television transmissions in Poland were started in the year 1937, using a system developed in the laboratories of the institute. After the second world war, the institute was re-established as the State Institute of Telecommunications, but in 1951 it was divided into several parts. A large part of the institute was moved into Miedzeszyn close to Warsaw (now a part of Warsaw) and named Instytut Łączności (official translation: National Institute of Telecommunications). This Institute served Polish economy and government by providing knowledge and expertise in the field of modern telecommunication technology. In the time of planned, closed economy in Poland, the Institute employed up to 1200 specialists; majority of them prepared Polish equivalents of international solutions of telecommunication devices. Such prototypes were subsequently transferred to be produced in Polish telecommunication industry and used by Polish telecommunication operator, or were used in co-operation with other countries of Middle and Eastern Europe.

In the last decade, National Institute of Telecommunications had to adapt itself to the opening of Polish economy. In the former, closed economy the Institute provided mostly technical expertise, with large part of tasks related to technical implementation and technology transfer or experimental production of telecommunication equipment. The opening of Polish economy resulted in an essential change of these tasks. In order not to loose the knowledge and expertise of the specialists employed by the Institute, we have chosen a difficult way of gradual changes of the character of the Institute. The difficulty related to the necessity of diminishing the size and changing the structure of employment, attracting young specialists, changing the priority to more basic research, and all the time preserving the economic viability of the Institute. In the beginning of the transformation, the long term tasks and goals of the Institute were not clear; but gradually become obvious that Poland needs a national research institution in such strategic field as modern telecommunications.

After the systemic transformation of Polish economy and society, gradual but rapid opening and liberalization occured first on the market for telecommunication equipment, later on the market for telecommunication services. The National Institute of Telecommunications had to change accordingly its orientation. But the deep knowledge of Polish telecommunication sector, resulting from the earlier role of the Institute, essentially helped in finding new directions. Thus, the Institute changed:

- from serving closed economy towards new international and European dimensions and co-operation;
- from concentration on applied technological research towards more basic research on new themes;
- in serving telecommunication industry – from prototype development towards testing quality and admissibility of telecommunication devices;
- from strictly disciplinary research in telecommunications towards interdisciplinary and integrated research including computer science and other information technology disciplines, law and economics of telecommunications, etc.;
- from strictly research and development orientation towards combining research with education activities, in particular – continuing and distant education;

- from concentration on telecommunication hardware towards increasing concentration on telecommunication software;
- from financing the Institute mainly from state subsidies towards financing the Institute mainly from market activities.

These changes were accompanied by an essential change in the size and structure of employment. In the five years of 1990–95, the number of employees decreased almost twice, to 670 in 1996. In the five years 1996–2000, the number decreased almost twice again, to 390 in 2000. At the same time, the Institute tried to employ $15 \div 20$ new young specialists yearly. Although labor markets in telecommunications are fluid and only 8 young specialists yearly remained for longer time at the Institute, the average age of specialists employed at the Institute decreased from 55 in 1996 to 49 in 2000[1] at the same time, the share of statutory public funds in financing of the Institute diminished to ca. 20%; the rest is obtained partly from competitive grants and mostly from market service.

During all these transformations, a new mission of the Institute was defined and formulated in 1999 as follows:

The mission of National Institute of Telecommunications is the development and application of knowledge and expertise – in the integrated fields of telecommunications, computer science (informatics) and other information technology sciences, including also legal, economic and social aspects of information technology – in the liberalised market of telecommunication and information services, according to the needs of emerging information society. Such knowledge and expertise should serve not only the operators and consumers on this market, but also governmental regulatory agencies responsible for this market. This is because the emergence of information society creates not only new and tremendous development chances, but also serious threats and dangers. For example, the development of new communication technologies creates large market chances, but also endangers correct market functioning, might violate the interests of consumers. Thus, it is necessary to understand well not only the developing technology, but also legal, economic and social issues related to its applications. A basic example of such relations are network interconnection issues in telecommunications. Therefore, an important element of the mission of the Institute is service with its knowledge and expertise to governmental regulatory agencies responsible for the functioning of telecommunication markets.

This mission – with small modifications – remains actual today. The programme of the Institute and the challenges it faces are strongly related to this mission. In order to understand better these challenges, we shall first discuss shortly more detailed results of structural changes of the Institute in last five years.

[1]This average is relatively low when compared with other research and development institutions in Poland, where a generation gap in science is developing.

2. Results of structural changes in National Institute of Telecommunications in five years 1996–2000

Structural changes had the objective to adapt the Institute to the fast changes in telecommunication technology and markets as well as to the liberalisation of telecommunication services. Several new units of the Institute had been formed, old units restructured or disbanded. New units include:

- Independent Section of Computerised Decision Support in Telecommunication;
- Independent Section of Theoretical Foundations of Telecommunications;
- Independent Section of Computer Networks, later transformed into Centre of Informatics;
- Division of Regulation and Economic Problems of Telecommunications;
- Division of Teletransmission and Fiber Optics;
- Division of Network Development and Applications of Informatics in Telecommunications;
- Independent Section of Marine Telecommunications and Radiocommunications in Gdańsk.

Restructured units include Centre of Scientific Information and Centre of Education and Promotion; the latter specialises in continuing postgraduate education for telecommunication engineers. Three small-scale production units – Division of Construction and Implementation in Warsaw and Production Division in Pułtusk as well as Production Section in Gdańsk were transformed – either in private enterprises or a smaller service units. At the same time, a Centre for Implementation of Information Technology and Services in Telecommunications specialised in implementations of computer technology, more in software than hardware.

At the same time, various administration and service units in the Institute were modernised, mostly with the help of a modern computer network organised in the Institute.

All these changes resulted in the following effects.

- The percentage of employees with university degrees was increased above 50%, the share of specialists with doctoral degrees and scientific titles[2] also increased significantly.

[2]In Polish scientific system there are two doctoral degrees and the term "scientific title" denotes the title of professor granted by the President of the Republic of Poland (corresponding roughly, but with higher prestige, to the position of full professor at Western universities).

- As the result of the increase of employed specialists with second doctoral degree or scientific title, the Scientific Council of the Institute regained its rights to grant doctoral degrees and exercised this right by starting to grant these degrees in the year 2000.
- The Institute obtained the highest category (1A) in the classification performed by the State Committee for Scientific Research.
- The Institute continued to organise the sequence of International Wroclaw Symposiums on Electromagnetic Compatibility and started to organise three new cyclical international conferences: International Experiences on Interconnection Issues, Research for Information Society, International Conference on Transparent Optical Networks. On the occasion of its 50-th anniversary, the Institute starts a new international conference – on Decision Support for Telecommunications and Information Society.
- In the year 2000, the Institute started to edit two new journals: one technical in Polish, *Telekomunikacja i Techniki Informacyjne* and one scientific in English, *Journal of Telecommunications and Information Technology.*
- Nine equipment-testing laboratories of the Institute obtained and maintained certified testing rights, issued by Polish Centre for Testing and Certification and by High Office of Measurements of Poland.
- The Centre for Education and Promotion of the Institute started postgraduate education in three specialities: the management of telecommunication networks, radiocommunication systems and multimedia telecommunication systems.
- The international co-operation of the Institute was strongly increased, including collaboration with many research institutions and international standard-setting organisations, with Framework Programmes of European Union, etc.

3. New tasks of the Institute

These results form a solid basis for further necessary changes of the Institute. A new law concerning research and development institutions introduces a new category of State Research Institute (SRI). Although full specification of this category is not available yet, the National Institute of Telecommunications aspires to this category and thus has to start several further changes in the years 2001–02. The most important of such changes are:

1) changes in the strategic research programme;
2) further changes in Polish and European co-operation, concentrating on the concept of centers and networks of excellence;
3) changes in the character of state services rendered by the Institute;
4) changes in the character of testing laboratories and certifications given by the Institute;
5) changes in the collaboration with diverse actors on telecommunication markets;
6) changes in the management system of the Institute, based on the concept of quality management.

Without describing all these directions in detail, we shall comment here only on the changes in the strategic research programme.

4. An outline of a strategic research programme

The preparation of a strategic research programme started with the discussions of the Scientific Council of the Institute that outlined several possible directions of future concentration of research, such as:

- the development of information society in its technical, legal and socio-economic aspects;
- terabit optical telecommunication networks;
- integration of diverse telecommunication and information technology systems together with their security aspects;
- planning and design of modern telecommunication networks;
- future trends of telecommunication and radiocommunication terminal devices;
- new techniques of telecommunication network management, in particular for IP protocol;
- new techniques of data analysis and computerised decision support in telecommunication and information services;
- new multimedia services, quality and standards of services, analysis of consumer demand on service markets;
- regulatory and economic aspects of Polish telecommunications, compatibility with European directives and standards.

Starting from these possible directions as an initial platform, several other considerations were included into discussions. These were:

- the ISTAG (*Information Society Technology Advisory Group* of European Commission) scenarios for ambient intelligence, prepared in co-operation with European Institute for Prospective Technologies in Seville in preparation of a new framework programme;

- an assessment of special chances of Polish telecommunication and information technology in the view of these European preparations;
- an assessment of the needs of forming information society in Poland and of the known programmes of scientific computer network infrastructure, such as *PIONER* of the State Committee for Scientific Research.

After such considerations, an outline of a strategic research programme was prepared, including a cooperation with several other research institutes in Poland. This outline consists of the following subprogrammes.

4.1. Design and management of secure and intelligent, integrated networks and systems with terabit capacities

The challenge in the development of modern communication networks is multimedia integration – not only of mobile and stationary telephony or data transmission, but also of internet and other communication services. The trend in stationary networks is towards fiber-optics and photonic technology, with migration towards IP protocol and an increase of transmission rates towards terabits per second. The technological development in this field will require costly research. In Poland, we must keep essential elements of knowledge in this field, but a chance and possible concentration of research should relate to necessary software development and the issues of *ambient intelligence* in such networks. Thus, especially interesting are the problems of design and management of such networks, together with the problems of security and quality of services. In these fields, it is possible to form a Polish research speciality. The goals of long-term research programme might include:

- methodology of design and planning of multimedia IP networks, together with their management systems;
- methods of monitoring and enhancing quality and security of services in multimedia IP networks;
- optics and photonic technology in terabit DWDM networks;
- wireless access systems to terabit backbone networks.

4.2. Decision support in telecommunications

The growing transmission rates, growing complexity of modern telecommunication networks, the diversity of services rendered and technologies applied, liberalisation of telecommunication markets and growing competition, all will increase the demand for computerised tools of decision support not only in management of telecommunication networks, but also their operators. This will concern diverse areas, such as acquiring and servicing clients as well as maintaining client relations, financial management, strategic management, negotiation of interconnection agreements, etc. This trend, visible internationally, will concern not only business actors on telecommunication markets, but also market regulation or electromagnetic spectrum management offices. We can take advantage of the fact that this is a relatively new field and that Poland has a strong, internationally known research school in decision theory and computerised decision support. Moreover, this field relates mostly to computer programming, hence the cost of research is lower than in technological fields. It is thus possible to make decision support in telecommunications a Polish specialty, concentrating e.g. on the following long-term research goals:

- data processing and analysis for decision support and the management of integrated multimedia networks;
- methods of decision support for the analysis of quality and security of integrated multimedia networks;
- decision support systems in monitoring electromagnetic spectrum use and the management of spectrum resources;
- methods of decision support in regulating interconnection issues.

4.3. Systems research and mathematical modelling for supporting e-business and e-banking

Electronic business and banking are today one of the most dynamic applications of integrated multimedia networks. Poland has good traditions of basic research in system and decision sciences as well as mathematical modelling; these tools are applicable to electronic business and banking. Thus, a cooperation of the National Institute of Telecommunications with such institutions as the Institute of Systems Research of Polish Academy of Sciences or the School of Applied Informatics and Management related to this Institute can result in the formation of a centre of excellence in this field. The research goals might include:

- mathematical models of electronic commerce transactions and their applications in the analysis of the development of electronic markets;
- systems of consumer support on electronic markets with application of software agents;
- mathematical models of electronic banking and their applications to the analysis of financial market development;
- systems of supporting investors in electronic banking, using multicriteria decision support and software agents.

4.4. Integrated teleinformatic systems for monitoring environment and supporting management of preventive actions in cases of natural impendences and disasters

Monitoring of diverse aspects of environmental pollution and hazards is subject of intensive research in Europe and in Poland, e.g. by such institutions as the Institute of Meteorology and Water Resources Management, or the Industrial Institute of Automatic Control and Measurements. Less developed is the integration of such monitoring systems. In cases of such natural disasters as deluges, it is necessary to integrate various sources of information in a multimedia network and to use them for supporting management and coordination of preventive actions. Such systems are being implemented, but they do not necessarily take into account the fast development of integrated multimedia telecommunication networks; thus, it is necessary to work on future generations of such systems. Possible research goals might include:

- review of diverse systems of monitoring natural environment, sources of information in such systems, and demand on diverse information services by various participants of preventive actions in cases of natural disasters;
- the development of a data model, use of data mining and analysis techniques and decision support methods in the management of preventive actions;
- the development of new concepts of integrated but dispersed teleinformatic systems for monitoring environment, with specific requirements for the security and quality of information services;
- the development and implementation of a pilot system on a regional level;
- the development of a prototype of a national coordination center based on next generation technology, together with requirements for data security, co-operation with preventive services in neighboring countries, etc.

4.5. Secure systems of information exchange for state and local administration, with special emphasis on information concerning labor market and distance learning

There are many technological variants, models and protocols of multimedia information exchange, including typical internet technologies, that can be used by central and local administration. An essential difficulty is related to the necessity of providing various levels of security in a network that must have public access for services such as electronic discussion groups, group review of proposed documents, electronic voting, etc. services needed in the concept of *e-democracy* on one hand, and provide high security in more special applications on the other hand. The development of cryptography, of diversified organisation of public and private keys, of electronic signature systems can provide necessary security levels even in networks with diversified access, but currently such techniques are expensive and are rapidly further developed – also by various institutions in Poland, e.g. by the National Academic Computer Network Institute. At the same time, the applications of such systems for information exchange concerning labor markets, or helping in distant, continuing education is an essential tool of fighting unemployment, with high priorities in European Framework Programmes. Thus, the challenge relates to the development of a system with diversified security requirements with applications to the integration of local information on labor market and on continuing education possibilities. This research should involve several institutions beside the National Institute of Telecommunications and the National Academic Computer Network Institute. The research goals might include:

- a review of multimedia information exchange techniques for an integrated system of supporting local and central administration, with elements of distance learning;
- a review of methods of providing graduated security levels for systems with diversified users;
- the development and pilot implementation of exchange of information about local labor markets and the possibilities of continuing education;
- the development and pilot implementation of an independent subsystem of continuing distant education with diversified levels and profiles of learning;
- the development and implementation of an education portal for the subsystem of distant education;
- implementation of such pilot systems for selected local administrations and for a selected central administration unit.

4.6. Multidisciplinary systemic analysis of social, economic, legal and technical aspects of information society and knowledge-based economy

The social, economic, legal and technical consequences and conditions of the development of information society and knowledge-based economy are a typical example of a complex multidisciplinary problem which should be addressed today by using the techniques and methodologies of systems analysis. The National Institute of Telecommunications started already preliminary research of such issues and organises since four years cyclic international conference *Research for Information Society.* However, there is a visible need of intensifying such research and searching for a systemic synthesis of its results. There might be several institutions involved in this research – beside National

Institute of Telecommunications, also Systems Research Institute of the Polish Academy of Sciences, Technical University of Warsaw, Main School of Commerce, Academy of Mining and Metallurgy in Cracow, etc. The goals of such research might include:

- the European Union before the challenges of information society and knowledge-based economy – a study of priorities, history, legal solutions;
- the conflicts of information civilisation era – the antithesis of uniformity of world information systems and the need of preserving cultural diversity, of information and work, of getting richer and increasing digital divide – and the means of overcoming them;
- an analysis of countries using intensively knowledge-based economy (such as Finland), their institutional and other means of promoting development;
- the role and character of educational systems and science systems in information civilisation and knowledge-based economy;
- the distinction of information and knowledge, types of knowledge representation, their standards, conditions and techniques stimulating knowledge exchange;
- the impact of telecommunications and information technology on information civilisation and knowledge-based economy; social and economic aspects of social acceptance of new technologies;
- the development of information society in rural areas; methods of neutralising digital divide.

5. Challenges before National Institute of Telecommunications at the turn of milenium

Many countries have national institutes working on telecommunications and information society technologies. Examples are *Federal Institute of Telecommunication Sciences* in Boulder, Colorado, USA, or *Royal Institute of Telecommunications* in Belgium. All such institutes have similar problems concerning financing research and attracting top-level researchers. Not all financing of such institutes comes from public funding, part comes from competitive grants or even directly from telecommunication market. The salaries of top-level specialists in this rapidly developing field are naturally much higher at telecommunication operators or equipment providing firms than at research institutes. There are diverse approaches to this problem; a typical one is treating the employment in a research institute as continuing education. Some of such institutes have majority of employees young but employed only for limited periods.

Thus, one challenge before the National Institute of Telecommunications is finding a way to implement such a recruitment model – where most young researchers would stay at the Institute only for a limited time. Such schooling of young employees is expensive, however, and would require a larger part (from 33% to even 50%) of the financing of the Institute provided from public funds. Two ways of reaching such proportion are possible. One is based on participating in long-term research programs, such as outlined above; the field of telecommunication and information society technology is certainly important enough strategically to deserve creating such long-term programs. Another way is further decrease of numbers and rejuvenation of the working force at the Institute. This is more difficult way, but it cannot be excluded, since typical institutes of this type in the world have no more than 200 employees.

The recruitment of new researchers might be intensified by broadening the educational activity of the Institute. The postgraduate studies and courses of continuing education, organised by the Education and Promotion Centre of the Institute, can be supplemented by a more close co-operation with diverse universities on engineer- and master-level education. In this, the idea of a center of excellence might help, stimulating the co-operation with educational institutions.

Another challenge relates to the growing interdisciplinarity of modern telecommunications and information society technologies. This challenge can be met by using systemic, multidisciplinary approach to research, as outlined in some of above sections.

The perspectives of the National Institute of Telecommunications, assuming that we can meet these challenges, are huge. The field of information society technologies for many decades yet will be decisive for the economic development of the world. We hope that after another fifty years, the National Institute of Telecommunications will celebrate the centennial anniversary of its activity.

Andrzej Piotr Wierzbicki – born June 29, 1937 in Warsaw. Graduated as Master of Engineering at the Faculty of Electronics, Warsaw University of Technology (WUT), in 1960. Ph.D. degree at this University in 1964, for a thesis on nonlinear feedback systems; D.Sc. degree in 1968, for a thesis on optimisation of dynamic systems. In 1971–75 a Deputy Director of the Institute of Automatic Control, later a Deputy Dean of the Faculty of Electronics, WUT. In 1975–78 the Dean of the Faculty of Electronics WUT. Since 1978 worked with the International Institute

for Applied Systems Analysis in Laxenburg n. Vienna, Austria; 1979–84 as the chairman of the theoretical branch, Systems and Decision Sciences Program, of this Institute. From 1985 back in the Institute of Automatic Control, WUT, as a Professor of optimisation and decision theory. In 1986–91 scientific secretary, currently member of presidium of the Committee of Future Studies "Poland 2000" (in 1990 renamed "Poland in XXI Century") of P.Ac.Sc. In 1991 elected a member of the State Committee for Scientific Research of Republic of Poland and the chairman of its Commission of Applied Research; contributed to basic reforms of Polish scientific system in 1991–94. Deputy chairman of the Council of Polish Foundation for Science in 1991–94, chairman of scientific councils of NASK (National Scientific and Academic Computer Network in Poland) and PIAP (the Industrial Institute of Measurements and Control). In 1991–96 the editor in chief of the quarterly "Archives of Control Sciences" of P.Ac.SC. In 1992 received (as first European researcher) the George Cantor Award of the International Society of Multiple Criteria Decision Making for his contributions to the theory of multiple criteria optimisation and decision support. Since 1996 the General Director of the National Institute of Telecommunications in Poland. In 2000 nominated as a member of the ISTAG (Information Society Technology Advisory Group) at European Commission. Since 2001 chairman of Advisory Group on Scientific International Cooparation of the State Committee for Scientific Research of Poland. Beside lecturing for over 40 years and promoting more than 80 master's theses at WUT (Warsaw University of Technology), he also lectured at the Department of Mathematics, Information Science and Mechanical Engineering of Warsaw University and in doctoral studies: at WUT, the Academy of Mining and Metallurgy, at the University of Minnesota, at the Illinois Technical University, Hagen University, and at the University of Kyoto. He also promoted 18 completed doctoral dissertations. Author of over 180 publications, including 11 books (4 monographs, 7 – editorship or co-authorship of international joint publications, over 50 articles in scientific journals (over 30 in international), 80 papers at conferences (68 at international, including over 48 published as chapters in books). He also authored 3 patents granted and applied industrially. Current interests include parallelisation of optimisation algorithms using multiple criteria approaches, diverse aspects of negotiation and decision support, including e.g. applications of fuzzy set theory for describing uncertainty in decision support models, multimedia software in computer networks, telematics in education, diverse issues of information society and civilisation. Languages: English, German, Russian (each fluent, beside native Polish). Member of IEEE, ISMCDM (International Society of Multiple Criteria Decision Making), SEP (Polish Society of Electrical Engineers), PTM (Polish Mathematical Society), PSKR (Polish Association for the Club of Rome).

e-mail: A.Wierzbicki@itl.waw.pl
National Institute of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland

Invited paper

Intra- and inter-cell interference investigations for broadband radio access systems above 10 GHz

Ranjan Bose, Andreas Hayn, and Rolf Jakoby

Abstract — Local multipoint distribution systems (LMDS) operating above 10 GHz have a large bandwidth (2–3 GHz) but a very limited range. These systems can provide coverage to a few kilometers only. The size of the macro cells illuminated by the base stations, where line of sight (LOS) exists, is 1 to 6 km. To provide coverage to customers where LOS is not possible, repeaters or passive reflectors may be used. In this paper we present first results of reflection measurements at 42 GHz, and based on that, simple multipath studies, taking into account the beamwidth of the antennas of both, transmitter and receiver. Secondly, LOS cochannel and adjacent channel interference are assessed for cellular LMDS networks. As suggested in the CRABS report, the maximal spectral efficiency can be obtained with a dual frequency and polarization reuse plan. This frequency and polarization reuse leads to interference. In this paper we have first calculated the cochannel interference (CCI) and the adjacent channel interference (ACI) due to the frequency/polarization reuse schemes suggested in the CRABS report. The effects of the variation of the half power beam width (HPBW) of the receiver, the time percentage parameter *p*, and the cell radius on *C/I* are also reported. In the latter part of the paper we propose a simple interference reduction technique based on re-orientation of the receiver antennas. We have also explored the possibility of using trellis coded modulation (TCM) for reducing interference levels. Initial results have been found to be quite encouraging.

Keywords — multipath propagation, interchannel interference.

1. Introduction

Broadband radio access systems operating at millimeter waves, used for local multipoint distribution service, have a large bandwidth of up to 2 GHz but a very limited coverage to a few kilometers only [1 ÷ 3]. This is mainly because they require clear line of sight between the base station and the subscriber antennas as well as because millimeter waves suffer on large propagation losses, particularly free-space propagation losses and attenuation caused by rain. However, in urban areas, the LOS requirements and the capacity needed for interactive services are more likely to limit the maximum cell size than the basic radio propagation characteristics [2, 3]. To enhance coverage to customers where LOS is not possible, passive and active repeaters may be used [4].

To compete successfully with the standard broadcasting, cable and satellites, the signals in the LMDS architecture need to be reliable and of high quality. The critical propagation issues are clear air absorption, signal attenuation by rain, vegetation and buildings, signal depolarization, multipath and cell-to-cell interference [5]. In Section 2 of this paper we investigate reflection and multipath effects and in Section 3 LOS cochannel and adjacent channel interference problems for broadband radio access systems above 10 GHz. The last section starts with the description of the LMDS system under study. Its first Subsection 3.1 deals with the LOS interference calculations. The effects of varying the receiver antenna beamwidth, the cell radius and the time percentage are reported in this section. In Subsection 3.2 we propose a simple technique for the reduction of cochannel interference by reorientation of the receiver antenna. Reorientation under the constraint of system availability is discussed in Subsection 3.3. The paper concludes by a discussion of the results in Section 4.

2. Reflection and multipath effects

Within LMDS cells we usually have a low penetration rate especially in urban areas. This is caused by the obstruction due to buildings, vegetation etc. An inexpensive option to increase penetration might be the usage of reflected waves into non-line-of-sight (NLOS) areas, if the reflection losses are smaller than a certain available system margin. This might be an option, especially in the vicinity of the hub below a radius of about 2 km where the system margin can be larger than 20 dB [6]. Besides the usage of such reflected waves, they always generate additional interference due to multipath propagation. For both, it is of interest to model the reflection losses realistically.

2.1. Reflection measurements

The reflecting wall or ground can be simply modeled in a first step by the Fresnel reflection coefficients for a plain dielectric surface and for perpendicular and parallel polarization:

$$R_{\parallel} = \frac{\sqrt{\underline{\varepsilon}_2 - \sin^2(\vartheta_{in})} - \underline{\varepsilon}_2 \cdot \cos(\vartheta_{in})}{\sqrt{\underline{\varepsilon}_2 - \sin^2(\vartheta_{in})} + \underline{\varepsilon}_2 \cdot \cos(\vartheta_{in})};$$

$$R_{\perp} = \frac{\cos(\vartheta_{in}) - \sqrt{\underline{\varepsilon}_2 - \sin^2(\vartheta_{in})}}{\cos(\vartheta_{in}) + \sqrt{\underline{\varepsilon}_2 - \sin^2(\vartheta_{in})}}, \qquad (1)$$

with the complex permittivity $\underline{\varepsilon}_2 = \varepsilon_2\,(1 - j\,\tan\,(\delta))$ of the dielectric material and the angle of incidence ϑ_{in} (Fig. 1). In a second step we assumed a slightly rough surface with Gaussian distributed roughness [8], having a standard deviation $\sigma_h < \lambda$. This leads to the following reduction coefficient [9]:

$$\rho(x) = e^{-x} \cdot I_0(x), \quad x = 8 \cdot \left(\frac{\pi \sigma_h \cos(\vartheta_{in})}{\lambda} \right)^2, \quad (2)$$

where I_0 is the modified Bessel function of the order zero and hence to the following modified Fresnel reflection coefficients:

$$R_{\|\text{rough}} = R_{\|} \cdot \rho\,; \quad R_{\perp\text{rough}} = R_{\perp} \cdot \rho\,, \quad (3)$$

which exhibits very good agreement with measurements of slightly rough and large surfaces compared to λ.

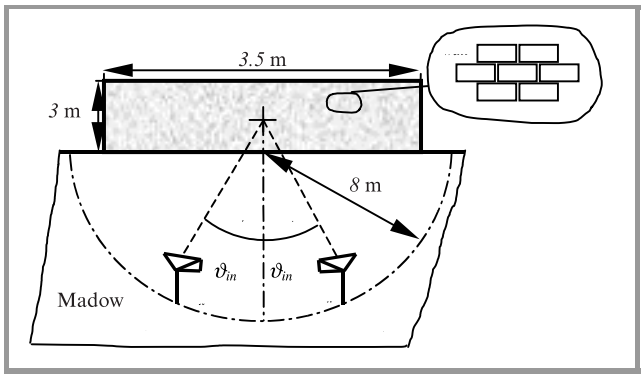


Fig. 1. Systematic reflection measurements on a brick wall.

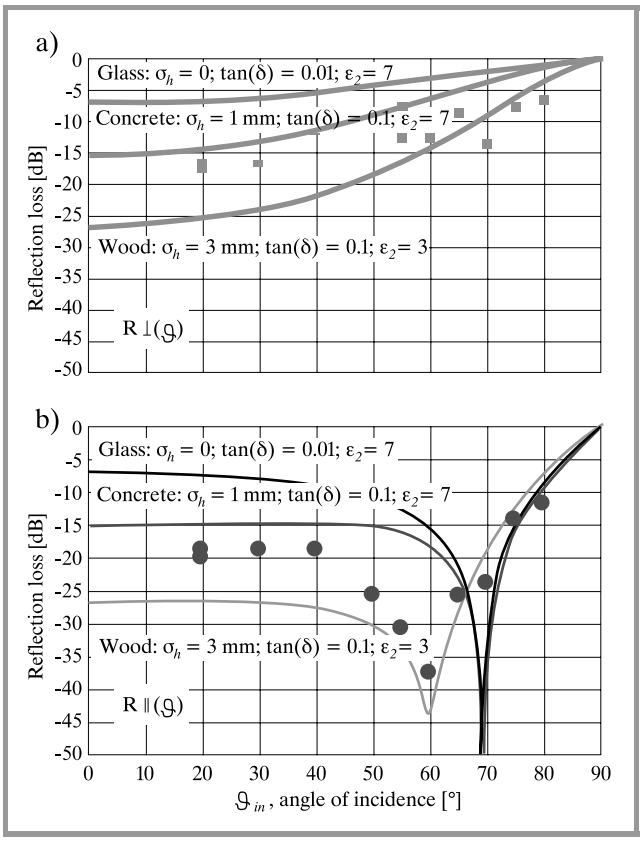


Fig. 2. Measured reflection loss of a brick wall and the modified Fresnel reflection losses for glass, concrete and wood for perpendicular polarization $R_{\perp}(\vartheta)$ (a) and parallel polarization $R_{\|}(\vartheta)$ (b) at 42 GHz.

For 42 GHz, Fig. 2 shows some results of systematic 2D-measurements of the bistatic reflection profile for $R_{\perp}(\vartheta)$ and $R_{\|}(\vartheta)$ from a brick wall with an average value of $\sigma_h \approx \lambda/7$, which was carried out by a measurement setup schematically shown in Fig. 1. Since the electric properties of the brick wall were not exactly known, we compared these measurements with the reflection losses calculated with Eq. (3) for three different materials: glass, concrete and wood. The measured reflection losses exhibits similar behavior vs. the angle of incidence as the calculated modified Fresnel reflection losses and are mostly in between the curves for concrete and wood, indicating of course, a large spread of about 5 dB due to some measurement uncertainties like slight misalignment of the very directional receiving antenna, temperature effects of the measurement equipment etc.

In order to model the reflection losses more realistically for LMDS scenarios, several systematic 3D-reflection measurements at 42 GHz were carried out at buildings in real environments with different materials, structures and surface roughness σ_h in dependence of the angle of incidence and along several measurement paths. One scenario is shown in Fig. 3a, where the LMDS transmitter was located at the top of a high building at 46.2 m and the received power of the reflections from a building front was recorded along the measurement path l at a height of 1.8 m above ground. The received power due to reflections from the building front along the path is indicated by the bars in Fig 3b. The variation in the received power of up to about 5 dB was partly caused by time variant effects like moving transmitter, receiver and vegetation being within the path due to wind forces as well as small alterations in the alignment of the receiving antenna around the direction where the maximum level was found.

To analyze 3D-reflections and to compare these measurements with prediction, two theoretical curves have always been plotted in addition: the free-space attenuation due to the wave travelling from the transmitter via the reflection point to the receiver (curve at the top) and this free-space attenuation plus the reflection loss (curve at the bottom). This last factor was calculated by firstly splitting up the incident field strength vector into parallel and perpendicular components to the plane of incidence, and then weighting both components with the corresponding modified Fresnel reflection coefficients for concrete according to Eq. (3). After that, the total field strength is given by summing up both vectors and hence the received power can be predicted by taking into account the polarization mismatch with the

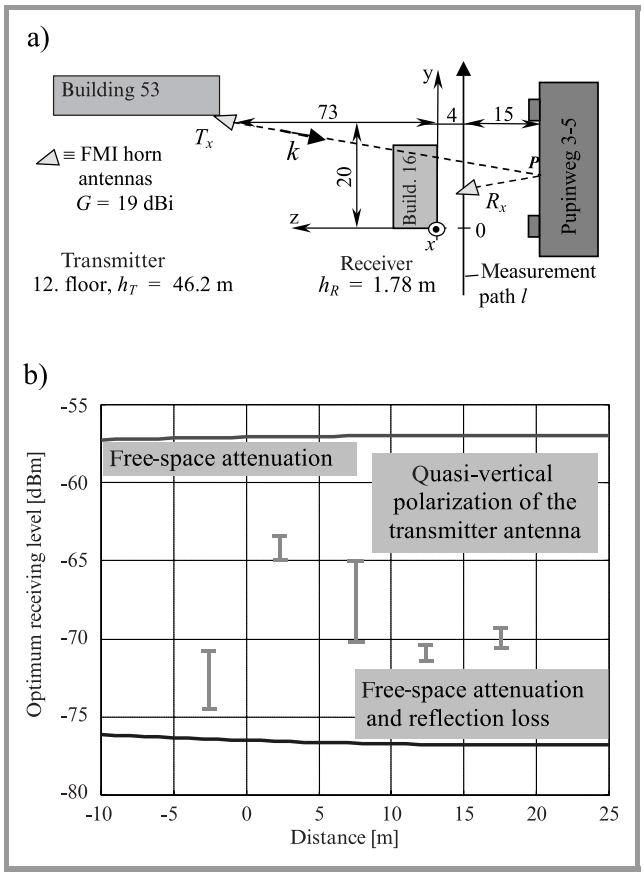


Fig. 3. (a) Typical LMDS scenario for measuring reflections from buildings; (b) received power of reflected waves recorded along a measurement path l in comparison with two theoretical curves.

receiving antenna, as well as the additional free-space attenuation. As we can see from Fig. 3b, all measured power levels (bars) are several dB above this predicted curve at the bottom, which has been valid for nearly all measurements. One reason for this higher measured levels than predicted for concrete ($\sigma_h = 1$ mm, $\tan(\delta) = 0.01$, $\varepsilon_2 = 7$) is of course, that all buildings consist of a mixture of different materials, in particular of concrete and glass, but also wood and metal frames. A second reason is that an exact alignment of the receiver antenna on the theoretical reflecting point of the buildings was very difficult. Therefore, we simply looked for the maximum of the received reflected power even it was found apart from the theoretical reflection line. This maximum is often caused by areas of the buildings with smooth window panes and metallic window frames, having relatively low reflection losses, and thus increasing the measured power levels compared to the predicted ones by assuming purely concrete with rough surfaces. Hence, these predicted curves for concrete walls are usually to pessimistic.

Beside that, our 2D- and 3D-measurements indicate, that on principle, the relatively strong reflections (except in the region of the Brewster angle for the parallel polarization) could be used to enhance the area coverage also into the NLOS areas when the system margin is sufficient. Thus, increasing the penetration rate of LMDS cells. Unfortunately, good reflectors of buildings are often windows, which are not reliable, because they can be opened and hence changing its scattering features. Hence, for using natural reflections, it has to be ensured that they are caused dominantly by static walls. A reliable option to enhance penetration into shadowed regions is the use of elliptical reflectors as passive repeaters [10].

2.2. Multipath propagation

Apart from using natural reflections to enhance penetration, strong reflections have always to be considered in conjunction with multipath propagation. This is generally agreed to be neglectable for LMDS, because of the very directional antennas of the customers, having half power beam widths usually in between 2° and 5°. But even, if the probability might be low to catch some multipath propagation components within this *HPBW*, there will partly be interference, if just one reflected wave in addition to the direct wave will be received within the *HPBW* of the customer antennas. Thus, it should be considered for the design of the equalizer of the receiver. Based on the reflection measurements and modeling those reflections for rough surfaces above, we carried out a worst-case estimation with a simple two-ray model according to the geometric arrangement in Fig. 4, assuming only one "optimally"-oriented slightly rough dielectric surface (dielectric reflector), which might be a building wall or a roof.

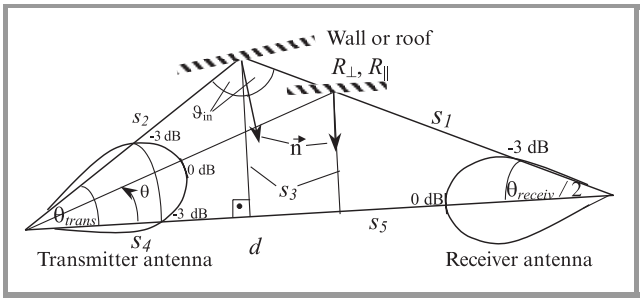


Fig. 4. Simple geometry for the estimation of worst-case multipath propagation (two-ray model).

For this simple arrangement, the direct ray is defined to go through the -3 dB point of the transmitter antenna and hits the maximum of the receiver antenna, whereas the reflected ray goes through an arbitrary point within the *HPBW* of the transmitter antenna under the angle θ, hits law of reflection and then cuts the receiver antenna at the -3 dB decay point. Hence, the antennas are sufficiently described by their *HPBW* θ_{trans} and θ_{receiv} and the antenna pattern of the transmitter within its *HPBW*, which is approximated by $\sin^2(\pi/4 + \theta)$ with $\theta \in [0, \pi/2]$ for the 90°-antenna. By the use of this simple geometrical arrangement, the magnitudes of both rays at the receiver, the relative time delay between the direct and the reflected waves as well as

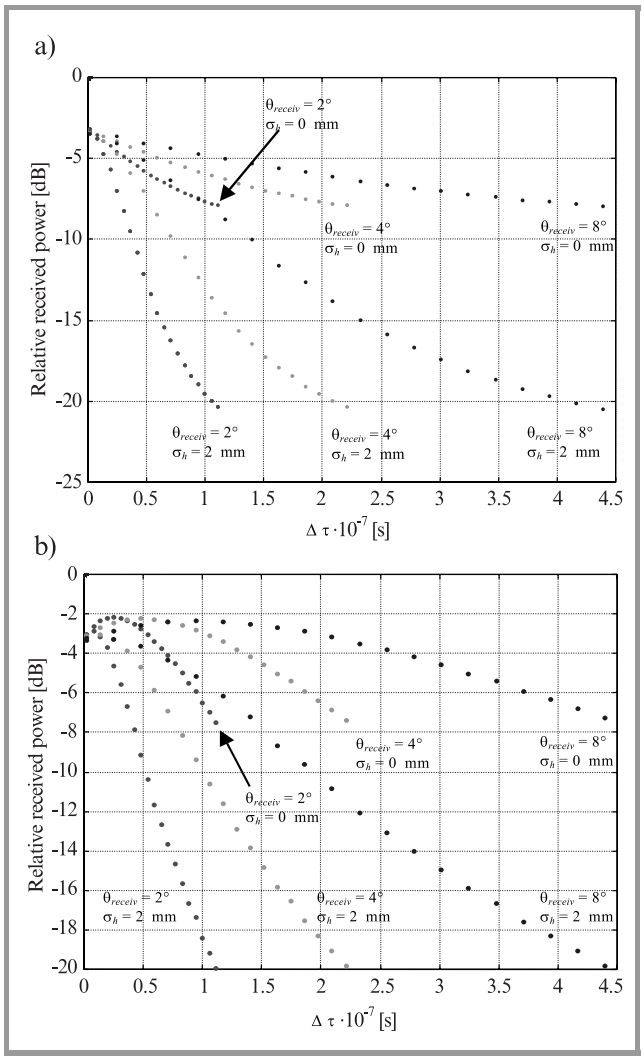


Fig. 5. Power delay profiles at the receiver with respect to the relative attenuation of the reflected path compared to the direct wave in dB with various roughness σ_h of the wall and *HPBW* Θ_{receiv} of the receiver antenna: (a) without considering a specific antenna pattern of the transmitter; (b) taking into account the $\sin^2(\pi/4+\theta)$ antenna pattern of the transmitter.

the power delay profiles were estimated for certain materials of the dielectric reflector in dependence of the distance d between the transmitter and receiver and the *HPBW*'s of both, the transmitter and receiver antennas. The relative time delay is:

$$\Delta\tau = \left(\frac{s_3}{\sin(\theta_{receiv}/2)} + \frac{s_3}{\sin(\theta)} - d\right)\cdot c^{-1},$$
$$\text{with} \quad s_3 = \left[\frac{1}{\tan(\theta_{receiv}/2)} + \frac{1}{\tan(\theta)}\right]\cdot d^{-1}. \qquad (4)$$

For typical LMDS receiver antennas with $\theta_{receiv} < 8°$, use can be made by the following linear approximation:

$$\Delta\tau \approx 0.32\cdot 10^{-3}\cdot\theta\cdot\theta_{receiv}\cdot d \quad [\text{ns}], \qquad (5)$$

where the parameters θ_{receiv} and θ are given in degrees and d in meter. The estimation error for this approximation remains low up to about $\theta_{receiv} = 16°$, where it exceeds 13%. For example, assuming $\theta_{trans} = 90°$, $\theta_{receiv} = 4°$ and a distance d of 2 km, the maximum relative delay $\Delta\tau_{max}$ can be 230 ns (i.e. if $\theta = \theta_{trans}$). So, even for narrow customer antenna pattern, this delay can be several times of the symbol rate of a QPSK-signal within a LMDS specific bandwidth of 33 MHz, which cannot be ignored for system design. It can easily be seen from Fig. 4 and Eq. (5), that each delay value $\Delta\tau$ has its corresponding value of the angle θ and hence, angle of incident on the reflecting surface:

$$\vartheta_{in} = \frac{1}{2}(180° - \theta_{receiv}/2 - \theta)\,. \qquad (6)$$

With this angle of incidence ϑ_{in}, the reflection losses can be estimated by the use of Eq. (3). As a worst-case estimation of a power delay profile, this has only been carried out for the perpendicular polarization, having less losses than the parallel polarization according to Fig. 2.

The envelope of the power delay profile can easily be derived with the close relationship between $\Delta\tau$ and this modified Fresnel reflection coefficient $R_\perp(\vartheta)$ due to the fixed geometrical arrangement according to Fig. 4. In a first iteration, it has not been taken into account a specific antenna characteristic of the transmitter antenna within $\theta_{trans} = 90°$. For the antenna configuration shown in Fig. 4, Fig. 5 shows the total received power at 42 GHz with respect to the one of the direct wave in dependence of the relative path delay $\Delta\tau$ for different θ_{receiv} and σ_h (e.g. glass and concrete), but fixed distance d = 2 km. It can be seen from Fig. 5a, that the envelope of the relative losses (i.e. power delay profile) is nearly proportional to the relative delay $\Delta\tau$ and θ_{receiv} of the receiver antenna. Taking into account the specified 90°-transmitter antenna characteristics and the other parameters as mentioned before, the power delay profiles in Fig. 5b are even worse. Comparing our results, the two ray models proposed in the IEEE 802.16 [11] standard are potentially to optimistic.

3. LOS interference

In this paper, we present the LOS calculations for the cell areas affected by the cochannel interference and the adjacent channel interference for a cellular LMDS architecture suggested by Telenor in the CRABS report [1, p. 46], which yields the same results as for a similar scheme proposed by Deutsche Telekom in [2]. In both, the maximal spectral efficiency was obtained with a dual frequency and polarization reuse plan according to Fig. 6, where the macrocells are square in shape. For this architecture, the cochannel cells appear in the 5th tier. The 42-GHz-LOS-interference calculations are done based on a H-plane or E-plane sectoral horn antennas for the transmitter at the base stations, providing horizontal and vertical polarization, respectively. The transmitter antennas have a half power beam width approximately equal to 90°. The receiver antennas of the subscribers are assumed to have a circular aperture with a parabolic taper on pedestal with a 10 dB

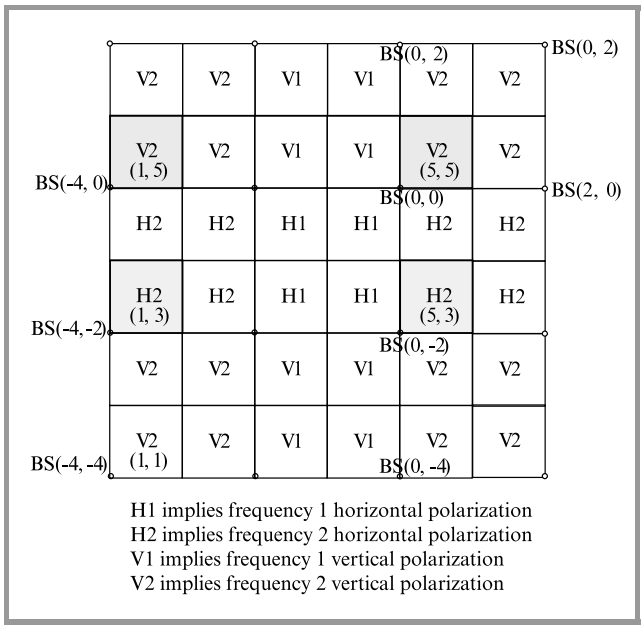


Fig. 6. Dual frequency and polarization reuse plan as suggested in the Telenor scheme of the CRABS report [1]. The circles in the figure denote the position of the base stations, on which four 90°-sectoral horn antennas are mounted. The two closest cells experiencing CCI are labeled (5, 5) and (1, 5) and those cells experiencing ACI are labeled (5, 3) and (1, 3), assuming the interfering cell is (1, 1).

edge illumination, having high gain between 30 to 40 dB with a very narrow *HPBW* of 5° to 2° for a diameter of about 10 to 24 cm.

The cochannel interference has been calculated for LOS under clear weather conditions without any power control strategy (worst case) using the formula:

$$C/I(L, L_2, p) = [EIRP - EIRP_1] + [G_R(0) - G_R(\theta)] + \\ -[L_{fs}(L) - L_{fs}(L_2)] - [A_{cs}(L) - A_{cs}(L_2)] + \\ -A_f(L_2, p) \quad [\text{dB}], \tag{7}$$

where $EIRP$ and $EIRP_I$ are the equivalent isotropic radiated powers of the desired and interfered signals at the customer location, $G_R(\theta)$ is the receiver antenna gain at an angle θ off the boresight, L_{fs} is the free space path loss, A_{cs} is the attenuation during clear sky, A_f is the short term enhancement due to atmospheric multipath and focussing effects, p is the time percentage for which A_f exceeds a certain value [2], L and L_2 are the distances in km from the base station of the desired and cochannel cell, respectively. At 42 GHz, A_{cs} has been taken as 0.2 dB/km [4, p. 92]. The short term enhancement due to atmospheric multipath and focussing effects, A_f, is determined using the formula [2]:

$$A_f = 2.6(1 - e^{-L/10}) \log_{10}(p/50). \tag{8}$$

3.1. LOS interference calculations

Numerical simulation was carried out to determine the percentage of areas where the C/I was above a specified threshold. Simulations were carried our both for cochannel and adjacent channel interference. The C/I levels of interest were 10, 15, 20, 25 and 3 dB, depending on the operating modulation scheme (QPSK, 8-PSK, 16-QAM or even 64-QAM).

The percentage of cell area below these C/I threshold levels are given in Table 1. The simulation takes into account a squared cell dimension $L = 2$ km, a time percentage $p = 0.01\%$ for which short term enhancement due to atmospheric multipath and focussing effects, A_f, exceeds 6.51 dB according to Eq. (8), and a receiving antenna of the customer with a diameter of 10 cm, i.e. an antenna gain of about 30 dB and a $HPBW$ of about 5°. The areas within the cell that experience these C/I levels (as a percentage of the total cell area) are shown in Fig. 7. Figure 7a gives the regions experiencing high CCI values while Fig. 7b depicts the regions with high ACI.

Table 1
Approximate areas within the cell that experience different levels of CCI

C/I value in cell (5, 5) [dB]	Approximate areas [%]	
	CCI	ACI
Below 10	5.48	0.14
Below 15	15.14	1.45
Below 20	22.24	4.80
Below 25	27.33	13.80
Below 30	29.87	24.85
Above 30	70.12	75.15

Next, simulations were carried out for different receiver antenna diameters/*HPBW* to see the effect of increasing the receiver antenna beamwidth. The results are given in Table 2. For these simulations, the transmitter is a H-plane sectoral horn and the receiver is a dish with a parabolic squared aperture. Interference from the most dominant cochannel cell has been considered only. The transmitter antenna characteristics are: $HPBW = 88°$, $A = 0.46$ cm ($0.65\ \lambda$), $R_1 = 2.00$ cm. For the simulation, the time percentage, $p = 0.01$ and the cell radius = 2 km.

In order to observe the effect of the time percentage, p, on the C/I, the parameter p was varied from 0.001% to 10%. The results are given in Table 3. It can be seen from the table that p, which has an effect on A_f (the short term enhancement due to multipath and focussing effects), has an important role to play in the determination of C/I. When we relax the time percentage parameter to even 0.1%, the C/I does not fall below 10 dB. However, p has a less severe effect on higher C/I. For example, a change in p from 0.001% to 10% causes the area where $C/I < 15$ dB to change by 89%, the area where $C/I < 20$ dB to change by 48% and the area where $C/I < 30$ dB to change by only 7%. Thus, the effect of p should not be neglected for low C/I.

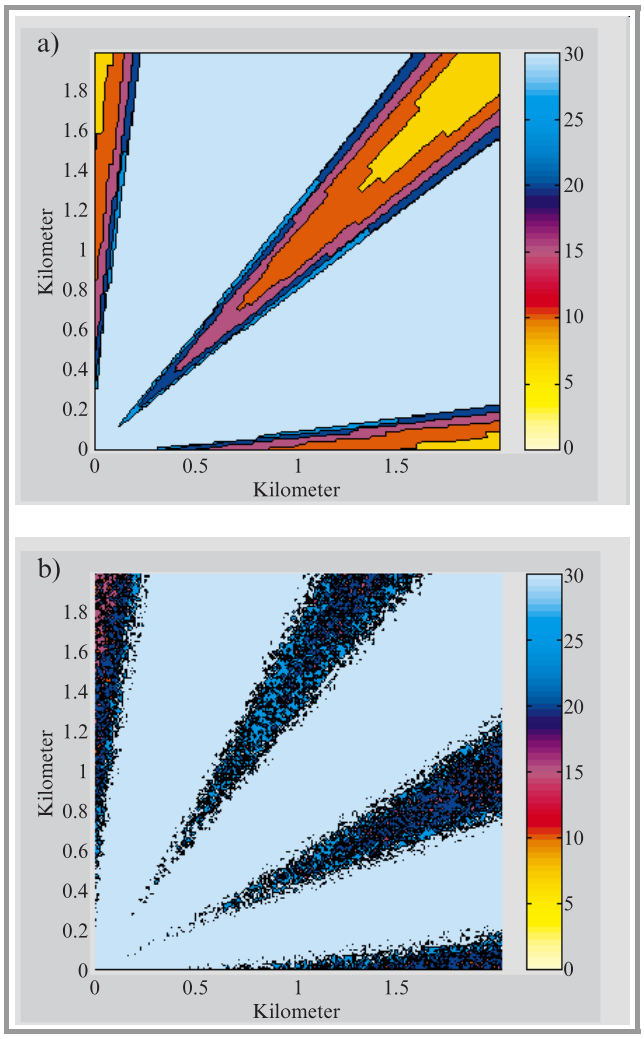


Fig. 7. Color coded areas within the cell (5, 5) that experience C/I levels below a specified level for (a) CCI and (b) ACI. The units of the color bar is in dB. The squared cell dimension is $L = 2$ km, the time percentage $p = 0.001\%$ for which A_f exceeds 7.81 dB and the *HPBW* of the receiver antenna is 5°.

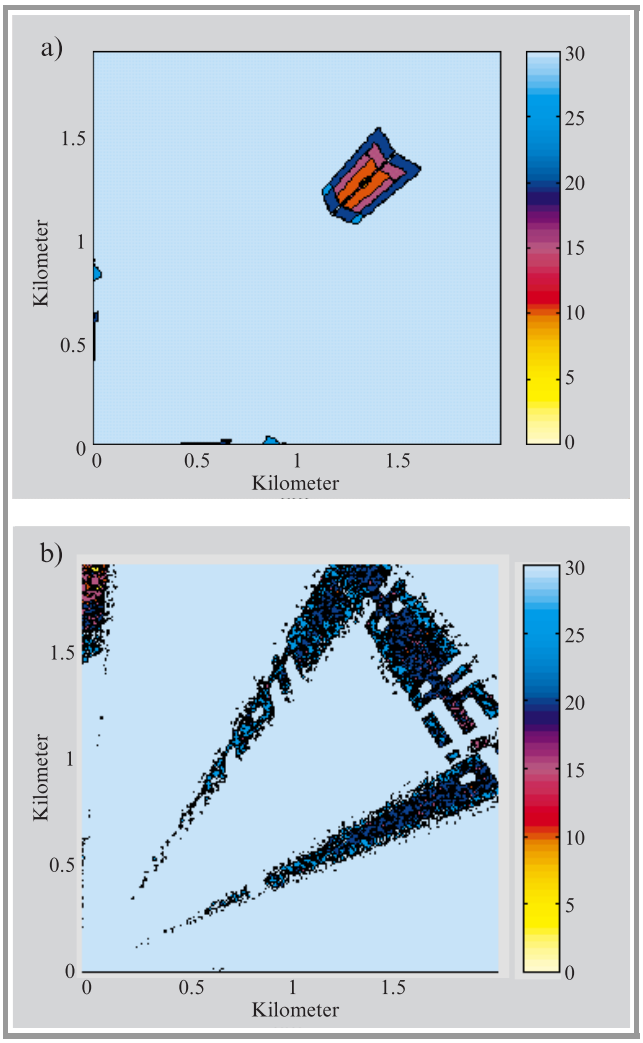


Fig. 9. The areas experiencing high levels of CCI (a) and ACI (b) in cell (5, 5) in the Telenor scheme after reorientation of receiver antennas. The units of the color bar is in dB. The *HPBW* of the $T_x = 88°$, that of the $R_x = 3°$, $p = 0.001$, cell radius = 2 km.

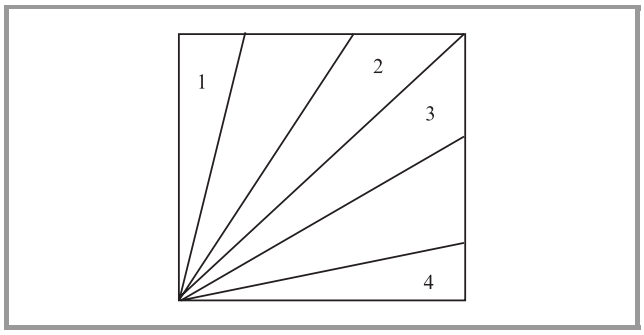


Fig. 8. The four disjoint wedges in which the receivers are reoriented.

Next, the effect of varying the cell radius on the interference was analyzed. The cell radius also plays a crucial role in the calculation of C/I. It has been mentioned earlier, the cell radius in the LMDS architecture can range from 1 to 10 km. Simulations were performed to find out the variation of C/I with the cell radius. The radius of the cell was varied from 0.5 to 15 km and the results are given in Table 4.

It can be observed from the table that the areas where the C/I is below a specified level at first increases with the increase in the cell radius. It reaches a maximum and then starts to decrease. The cell radius that results in the minimum C/I (i.e., the maximum areas experiencing high interference) is approximately 3 km, for the simulation parameters considered here. This behavior can be explained as follows. At very low values of the cell radius, the free space path loss, A_{fs}, is the dominant factor. As the cell radius increases, the short term enhancement due to multipath and focussing effects, A_f, starts playing a dominant role and increases the interference level. As we increase the cell radius further, the increase in the interference due to A_f is compensated by the free space path loss, A_{fs}. Consequently, the total interference starts decreasing.

Table 2
Percentage of areas below a specified C/I ratio in cell (5, 5)

R_x antenna		Cell (5, 5)				
Diameter [cm]	*HPBW* [°]	$C/I < 10$ dB	$C/I < 15$ dB	$C/I < 20$ dB	$C/I < 25$ dB	$C/I < 30$ dB
24	2.0	0.15	1.08	1.52	1.95	2.51
20	2.4 (CRABS)	0.82	4.44	7.22	8.93	9.40
10	4.7	1.94	8.52	13.08	15.59	18.41
9.6	5.0	2.00	8.99	13.85	16.97	19.05

Table 3
Percentage of areas below a specified C/I ratio in cell (5, 5) (diameter of the receiver = 20 cm, $HPBW = 2.4°$)

Time percent, p [%]	$C/I < 10$ dB	$C/I < 15$ dB	$C/I < 20$ dB	$C/I < 25$ dB	$C/I < 30$ dB
0.001	2.68	4.99	8.12	9.17	9.47
0.01	0.61	4.23	6.85	8.83	9.38
0.1	0	3.55	5.65	8.53	9.29
1.0	0	1.92	4.84	7.86	9.12
10	0	0.51	4.20	6.71	8.84

Table 4
Percentage of areas below a specified C/I ratio in cell (5, 5) (diameter = 20 cm, $HPBW = 2.4°$, $p = 0.01\%$)

Cell radius [km]	$C/I < 10$ dB	$C/I < 15$ dB	$C/I < 20$ dB	$C/I < 25$ dB	$C/I < 30$ dB
0.5	0	2.00	4.88	7.92	8.88
1.0	0	3.54	5.49	8.48	9.25
2.0	0.61	4.23	6.85	8.83	9.38
3.0	0.71	4.36	7.06	8.89	9.40
5.0	0.28	3.95	6.29	8.74	9.36
10.0	0	0.51	4.22	6.75	8.87
15.0	0	0	0.59	4.32	6.92

3.2. Reducing CCI using reorientation of receiver antennas

Having investigated the LOS cochannel and adjacent channel interference problem, we now propose a simple technique for the reduction of areas within the cell experiencing high levels of interference. Let us assume that the current BS location is (0, 0) as given in Fig. 1. It should be noted that the coordinate system for the BS is different from that used for representing the cells. We need two different coordinate systems because the BS are not placed at the center of the cell, but at one of the corners. Hence, in many cases the same BS location is valid for four cells. From Fig. 7 we note that the regions of high interference roughly form wedges within the cell. There are three disjoint wedges where the interference levels are unacceptable due to the three cochannel cells: (1, 5), (1, 1) and (5, 1). The receivers in these regions of high interference can be reoriented to the BS of nearby cells in order to reduce interference levels. Here we have taken the cut-off $C/I = 30$ dB to label an area as a high interference zone. By simply observing the cells around the wedges that depict high interference regions, we choose the reorientation scheme as following. The receivers located in wedge 1 are reoriented to BS at (2, 2), those in the top section of wedge 2 to BS at (0, 2),

Table 5
After reorientation, the current and the interfering BS for the four wedges

Wedge	Current BS (after reorientation)	Interfering BS (CCI)		
		dominant	less dominant	
1	(2, 2)	(6, 6)	(10, 6)	–
2	(0, 2)	(−4, 10)	(−4, 6)	(−4, 2)
3	(2, 0)	(10, −4)	(6, −4)	(2, −4)
4	(2, 2)	(6, 6)	(6, 10)	–

those at the bottom half of wedge 2 to BS at (2, 0) and those in wedge 4 to BS at (2, 2). Thus, there are now four wedges from the point of view of reorientation, as depicted in Fig. 8. Due to reorientation, new cochannel cells come into picture. Table 5 gives the list of the new cochannel cells for the four wedges after reorientation.

As a first level interference estimation, only the dominant cochannel cell was considered for line of sight CCI calculations after reorientation. The areas facing high interference after reorientation is shown in Fig. 9. Table 6 compares the percentage of areas suffering from high CCI values before and after the reorientation scheme. If, upon reorientation of a receiver antenna the CCI worsens, it is reverted back to the original base station.

Since we make the best choice for a receiver regarding which BS it should point to, the reorientation strategy can only provide improvement. It can be seen in Fig. 9 that there are still some areas where, even after reorientation, the interference level is unacceptable. However, the levels are better than before. For example, the worst C/I value prior to reorientation was 6.65 dB, and after reorientation, the worst value is $C/I = 9.74$ dB (over 3 dB improvement, though still unacceptable).

We next look at the effect of reorientation of receiver antennas on adjacent channel interference. The process of reorientation to reduce CCI also results in new ACI cells. Table 7 lists the new cells that offer ACI as a result of reorientation in order to reduce the problem of CCI. The values of the resulting ACI due to the reorientation of R_x antennas is given in Table 8. The plot of the areas experiencing ACI is also shown in Fig. 9. From Table 3 it is clear that the process of reorientation has not reduced the ACI much. The values are comparable to those obtained prior to reorientation. This was expected as the strategy for reorientation of R_x antennas was to reduce CCI, and not ACI. It could be possible that there is a slight increase in the ACI values after reorientation.

Table 6
Approximate areas within the cell that experience different levels of CCI before and after reorientation of R_x antennas

C/I value [dB]	Approximate areas [%]	
	before reorientation	after reorientation
Below 10	3.52	0.04
Below 15	9.71	0.70
Below 20	13.05	1.56
Below 25	17.05	3.12
Below 30	19.12	3.47
Above 30	80.87	96.52

A more appropriate solution would be to check every R_x location and see whether the reorientation to reduce CCI or ACI produces a better result. The reorientation should be done accordingly. In many cases, one of the two, CCI or ACI increases and the other one decreases due to reorientation. Plus a third choice is not to go for reorientation. So the strategy should be to choose the best of the three possible solutions using the following decision rule: "Choose orientation to maximize the minimum of (CCI_i, ACI_i) for the i^{th} choice" i.e.,

$$\text{decision variable} \quad d_i = max\left(min\left(CCI_i,\ ACI_i\right)\right). \qquad (9)$$

Table 7
After reorientation, the current and the interfering BS for ACI calculations

Wedge	Current BS (after reorientation)	Interfering BS (ACI)		
		dominant	less dominant	
1	(2, 2)	(6, 4)	–	–
2	(0, 2)	(−4, 8)	(−4, 12)	–
3	(2, 0)	(6, −2)	(10, −2)	(10, −6)
4	(2, 2)	(6, 4)	–	–

Table 8
Approximate areas within the cell that experience different levels of ACI before and after reorientation of R_x antennas

C/I value [dB]	Approximate areas [%]	
	before reorientation	after reorientation
Below 10	0.10	0.17
Below 15	0.83	1.18
Below 20	2.71	3.43
Below 25	8.08	9.07
Below 30	14.66	15.55
Above 30	85.33	84.44

3.3. Reorientation under the constraint of system availability

In actual scenarios, reorientation of receiver antennas at all locations is not possible because of the system availability

Table 9
Approximate areas [%] within the cell that experience different levels of interference after the reorientation of R_x antennas, with and without system availability constraint ($p = 0.01\%$ and receiver $HPBW = 2.4^\circ$)

C/I value [dB]	CCI		ACI	
	without constr.	99.7% sys. avail.	without constr.	99.7% sys. avail.
Below 10	0	0.16	0.03	0.00
Below 15	0.51	2.35	0.50	0.42
Below 20	0.52	3.46	1.84	1.57
Below 25	1.54	5.71	5.46	5.06
Below 30	1.58	6.04	10.88	10.18
Above 30	98.42	93.95	89.11	89.81

constraint. If we try to orient the receiver antennas to far off alternate base stations, the LOS may not exist due to tall buildings and foliage. To get a feel for the performance due to the system availability constraint, we put in the following restriction: we do not reorient those receiver antennas for which the alternate base station locations are more than 1.5 times the cell radius (one side of our square cell). This factor, r, equal to 1.5 approximately corresponds to 99.7% system availability. The results of the CCI and ACI with this constraint is given in Table 9 for $p = 0.01\%$.

Table 10
Approximate areas [%] within the cell that experience different levels of interference before reorientation of receiver antennas

R_x *HPBW*	C/I value [dB]	$p = 0.01$		$p = 0.001$	
		CCI	ACI	CCI	ACI
2.4°	Below 10	0.79	0.03	2.60	0.07
	Below 15	6.13	0.40	6.95	0.64
	Below 20	9.89	1.56	11.53	2.15
	Below 25	13.15	4.68	13.57	6.43
	Below 30	14.01	9.80	14.11	11.36
	Above 30	85.98	90.19	85.89	88.64
5.0°	Below 10	1.61	0.05	5.48	0.16
	Below 15	12.30	0.85	15.14	1.38
	Below 20	20.40	3.35	22.24	4.63
	Below 25	25.24	9.84	27.33	13.44
	Below 30	29.34	20.76	29.87	24.58
	Above 30	70.65	79.24	70.12	75.41

Table 11
Approximate areas [%] within the cell that experience different levels of interference after reorientation of R_x antennas

R_x *HPBW*	C/I value [dB]	$p = 0.01$		$p = 0.001$	
		CCI	ACI	CCI	ACI
2.4°	Below 10	0.16	0.00	0.43	0.05
	Below 15	2.35	0.42	2.57	0.74
	Below 20	3.46	1.57	5.02	2.30
	Below 25	5.71	5.06	5.87	7.22
	Below 30	6.04	10.18	6.08	11.73
	Above 30	93.95	89.81	93.92	88.26
5.0°	Below 10	0.23	0	1.42	0.12
	Below 15	5.32	0.72	7.70	1.45
	Below 20	10.60	3.01	12.39	4.50
	Below 25	13.45	9.51	16.34	13.30
	Below 30	17.59	19.58	17.80	22.97
	Above 30	82.40	80.41	82.19	77.02

It can be observed that, under this constraint, there is a decrease in the system performance with respect to the CCI, as expected. However, there is an improvement in the levels of ACI. This is because, in the first case where we reorient with the sole objective to reducing CCI, in many cases we end up increasing the ACI in the process. However, when we reorient under the system availability constraint, we only reorient some of the receiver antennas. In order to get a feel for the typical and the worst case scenarios, we obtain the CCI and ACI for receiver *HPBW* of 2.4° (typical) and 5° (worst case). We also carried out simulations for $p = 0.01$ (typical) and $p = 0.001$ (strict). The interference levels before and after reorientation are tabulated in Table 10 and Table 11 respectively.

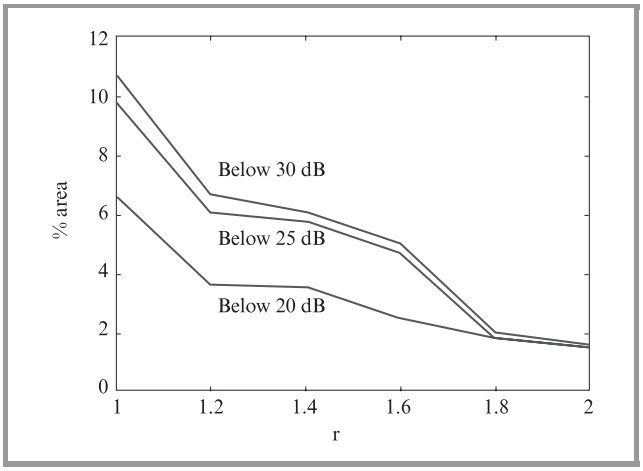


Fig. 10. The percentage of areas in the cell (after reorientation of receiver antennas) versus the multiplicative factor r which determines the system availability.

The extra restriction of the system availability constraint results in a degradation in the system performance, as indicated earlier. In Fig. 10 we plot the percentage of areas in the cell (after reorientation of receiver antennas) versus the multiplicative factor r which determines the system availability. For example, $r = 1.5$ implies that receiver antennas will be reoriented to an alternate BS if the LOS distance does not exceed 1.5 times the cell radius. This corresponds to system availability of 99.7%.

4. Conclusions

In this paper we have investigated the intra-cell interference as well as LOS cochannel and adjacent channel interference for the LMDS architecture. Simulations have beenperformed at 42 GHz. Our reflection measurements and multipath propagation studies have shown, that even for very directional customer antennas ($HPBW = 2°$) strong mulipath components might occur and that they are not neglectable. Therefore, the multipath propagation has to be considered in the design of the equalizer of the receiver.

It has been observed that about $10 \div 15\%$ of the areas suffer from interference problems (for an acceptable threshold

of $C/I > 20$ dB) under LOS and clear weather conditions when cell dimensions of about 2 km are deployed. An interesting thing that has been observed is that for a fixed percentage p, the percentage of cell area below these C/I threshold levels firstly increases with the cell dimension L, reaching its peak value at about 3 km, and then decreases again. The LOS interference problem will be reduced if the *HPBW* of the subscriber's antennas are decreased down to 2°, which means comparatively large antenna diameters of up to 25 cm and more difficulties of arranging them to point precisely to the base station, thus suffering some alignment losses.

In the latter portion of the paper we have proposed a simple technique to reduce the cochannel interference simply by reorienting the receiver antennas to a more favourable base station. It was observed that by reorientation, the area encountering CCI below 30 dB was decreased from 19.1% to 3.4%. This big improvement comes at no additional cost. The ACI values remain comparable to those prior to reorientation. If the system availability constraint is put, we still observe a big improvement, though not as large as without any constraint. Further improvement is possible if we check every receiver location and see whether the reorientation to reduce CCI or ACI produces a better result.

References

[1] CTS Project 215, "Cellular Radio Access for Broadband Services (CRABS)", Feb. 1999, pp. 18–19.

[2] ITU-R Rec. 452-6, "Prediction Procedure for the Evaluation of Microwave Interferrence between Stations on the surface of the Earth at Frequencies above 0.7 GHz", PN Series, Propagation in Non Ionized Media, 1990, pp. 565–591.

[3] A. Seville, M. Willis, and E. Falaise, "Area coverage studies for millimetre-wave services", in *Millennium Conf. Anten. Propag.*, Davos, Switzerland, Apr. 2000.

[4] R. Jakoby and M. Grigat, "MMDS zur Erweiterung von BK-Netzen", Research Report from the Research Center of Deutsche Telekom AG, 1996.

[5] P. B. Papazian, G. A. Hufford, R. J. Achatz, and R. Hoffman, "Study of the local multipoint distribution service radio channel", *IEEE Trans. Broadcast.*, vol. 43, no. 2, pp. 1–10, 1997.

[6] A. Hayn and R. Jakoby, "Propagation and standardization issues for 42 GHz digital microwave video distribution system (MVDS)", in *Tech. Rep. COST 259 TD (99) 021*, Thessaloniki, Greece, Jan. 1999.

[7] A. Hayn, G. Bauer, J. Freese, and R. Jakoby, "Entwicklung von Ausbreitungsmodellen und Software-modulen zur Abschätzung der MWS-Versorgung unter Berücksichtigung von abgeschatteten Bereichen", Research Report for T-Nova, Deutsche Telekom, Nov. 1999.

[8] W. S. Ament, "Toward a theory of reflection by a rough surface", *Proc. IRE*, vol. 41, no. 1, pp. 142–146, 1953.

[9] L. Boithias, *Radio Wave Propagation*, New York: McGraw-Hill, 1987.

[10] G. Bauer, J. Freese, and R. Jakoby, "Single-cell coverage prediction of LMDS including passive reflectors", in *ICAP 2001*, Manchester, Apr. 2001.

[11] IEEE 802.16.1pc-00/15, "Proposed System Impairment Models", Feb. 2000.

Ranjan Bose received his B.Tech. in electrical engineering from IIT, Kanpur in 1988 and his M.Sc. and Ph.D. in electrical engineering from the University of Pennsylvania, PA in 1993 and 1995, respectively. He worked as a Senior Design Engineer in Alliance Semiconductors Inc., San Jose, CA from 1996 to 1997. Since November 1997 he is working as an Assistant Professor at IIT, Delhi in the Department of Electrical Engineering. His research interests lie in the areas of broadband wireless access and coding theory. He received the URSI Young Scientist award in 1999 and is currently at the Technical University of Darmstadt, Germany on the Humboldt Fellowship.
e-mail: rbose@ee.iitd.ernet.in
Technische Universität Darmstadt
Institut für Hochfrequenztechnik
Fachgebiet Mikrowellentechnik Merckstrasse 25
D-64283 Darmstadt, Germany

Andreas Hayn received the Dipl.-Ing. (M.Sc.) degree from the Technical University of Darmstadt, Germany in 1997. He is currently working towards the Dr.-Ing. (Ph.D.) degree in electrical engineering at the Institute of Microwave Engineering at the Technical University of Darmstadt. His research interests include the electromagnetic and acoustic wave propagation and broadband wireless radio systems.
e-mail: a.hayn@ieee.org
Technische Universität Darmstadt
Institut für Hochfrequenztechnik
Fachgebiet Mikrowellentechnik Merckstrasse 25
D-64283 Darmstadt, Germany

Rolf Jakoby received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from the University of Siegen, Germany, in 1985 and 1990, respectively. In 1991, he joined the Research Center of Deutsche Telekom in Darmstadt, Germany. There, he has first conducted radiowave propagation research in the Ku- and Ka-band within ESA's

OLYMPUS satellite campaign and in the UHF-band for UMTS mobile communications in urban microcells. Then, he lead a project concerning Local Multipoint Distribution Systems (LMDS) at 42 GHz. In April 1997, he got a professorship for Microwave Technology at the Technical University of Darmstadt, where his research is focused on broadband wireless systems and on adaptive antennas at microwaves. He is a member of the IEEE and the German Society for Information Technology ITG. In Feb. 1992, he received a prize and award for his Ph.D. thesis from the CCI Siegen and in Dec. 1997, a prize and award from the ITG for an excellent scientific publication in the IEEE AP-44, 1996.

e-mail: jakoby@hrzpub.tu-darmstadt.de
Technische Universität Darmstadt
Institut für Hochfrequenztechnik
Fachgebiet Mikrowellentechnik Merckstrasse 25
D-64283 Darmstadt, Germany

Invited paper

Radio wave propagation and single-cell coverage prediction of broadband radio access systems including passive reflectors

Georg Bauer, Jens Freese, and Rolf Jakoby

Abstract — This paper presents the scheme of a software planing tool for single microcells of broadband radio access systems operating at millimeter waves, providing local multipoint distribution services (LMDS), particularly in urban or suburban areas. The aim of this planing tool is to provide operators with information, which supports the assessment of the profitability by calculating link budgets and the area coverage and by roughly estimating the maximum number of potential customers, i.e. the total number of households in a certain area, as well as the number of customers (households) which will actually be covered from a site. A household is here considered as a single apartment or flat, thus a building usually consists of several households. This novel household-estimation model is able to estimate the number of households based only on the 3D-data of buildings, without using residential data. Besides the optimization of line-of-sight (LOS) coverage, also the possibility is considered to enhance the coverage into shadowed areas by using optimized reflectors as passive repeaters up to distances of 1 km from the base station, depending on the system margin and the rain climate zone. All calculations are based on 3D-databases of both, buildings and vegetation.

Keywords — millimeter wave propagation, LMDS, coverage prediction, passive reflectors.

1. Introduction

Broadband radio access systems operating at millimeter waves, used for local multipoint distribution services, are local cellular point-to-multipoint radio systems, delivering multimedia services and/or broadcast services from a central transmitter or base station to individual houses or blocks of apartments or houses for residential and business customers within its cell size, offering rapid infrastructure deployment and the ability to provide local content. It can be significantly cheaper to install than a cable system since only homes requesting LMDS are provided with receivers (extension on demand). Hence, these LMDS systems can be applied as a supplement of the broadband cable TV network (CATV), as a substitution of the CATV or direct broadcast via satellite (DBS) or as a cable pull through during the building up or extension of parts of the cable networks. For these reasons and because of its flexible installation, LMDS systems have recently gained popularity worldwide, in particular in the USA, Canada and in Eastern Europe.

Figure 1 exhibits an example of a bi-directional cellular LMDS system primarily for broadcast services to residential customers. The different sources, e.g. from satellite, terrestrial VHF/UHF or local contents are collected in a Head End, and then provided to the LMDS base stations via a backbone network, which might consists of radio links in the millimeter wave region, optical fiber or other cable networks. Beyond one-way broadcasting distribution, this wireless technology has the potential 1) for providing expanded TV services like pay per view, near or video on demand as well as, 2) to become a low-cost integrated medium for providing voice and data services, as well, having the potential to integrate the CATV and ISDN networks. This allows fully interactive services such as video conferences and high speed access to Internet. This wireless technology can be a quick, simple and cost-effective solution.

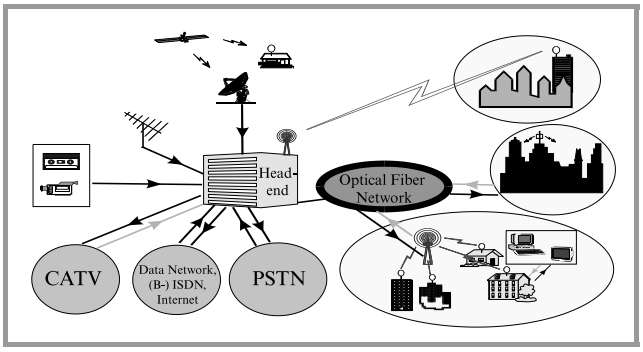


Fig. 1. Example of a bi-directional cellular LMDS system.

Local multipoint distribution services systems at millimeter waves have a large bandwidth of up to 2 GHz but a very limited coverage to a few kilometers only [see Section 3]. This is mainly because they require clear line of sight for reliable point-to-multipoint links between the base station and the subscriber antennas as well as because millimeter waves for LMDS in between 26 GHz and 43 GHz suffer on large propagation losses, particularly free-space propagation losses and attenuation caused by rain as well as large diffraction and reflection losses on buildings.

Hence, in literature, very often only the number of covered buildings within a cell is used to estimate the penetra-

tion rate of a LMDS microcell. It is usually calculated by checking only the roof tops of the buildings for coverage. In general, however, the number of covered buildings does not agree with the number of customers, since there are often several households (customers) in one building, which might be served only via one customer station located on the roof top. However, similar to satellite broadcasting, many customers, in particular residential subscribers, might wish to have direct access to a broadband multimedia wireless network. As a consequence, the percentage of customers that can be covered by one LMDS base station is much lower than the number of covered roof tops. While a roof top coverage (buildings) of up to 70% can usually be achieved in a area range of 1 km around the base station Biddiscombe [3], only $25 \div 35\%$ of the customers (households) can be supplied with a direct access. Even near the base station there are a large number of uncovered households. Therefore, a simple household-estimation model is presented to get more realistic results, and in addition, the use of a shaped metallic reflector is introduced in order to enhance the area coverage into shadowed areas. Both, the household-estimation model and some approximated results of the reflector analysis with physical optics (PO) Diaz [4] have been implemented into a software tool for single-cell LMDS coverage prediction.

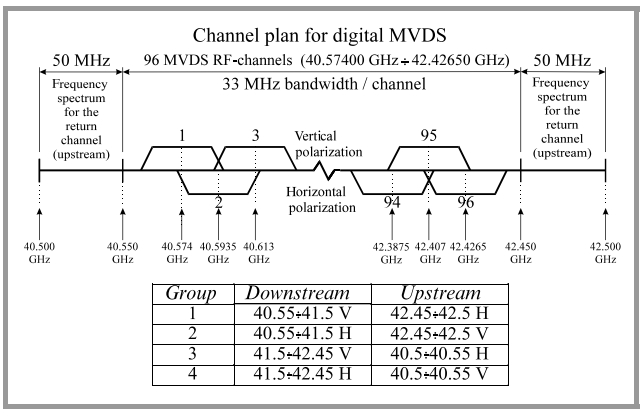
2. Example of an unidirectional LMDS system for broadcast services

There are many standardization boards and hence different standards for LMDS worldwide. As a standardization example in Europe, the $40.5 \div 42.5$ GHz band has been harmonized within the CEPT for a multipoint video distribution system (MVDS) or unidirectional LMDS system. The standards for digital 42 GHz MVDS [3] are based on performance specifications from working groups in the UK [7] and the DVB-Adhoc group [2, 8]. Hereafter, the system performs the adaptation of the base band TV signals from the output of the MPEG-2 transport multiplexer to the MVDS channel characteristics. Then, it uses randomization for energy dispersal and a concatenated error protection strategy based on a shortened Reed-Solomon (RS) code with a code rate $r_1 = 188/204 = 0.922$, convolution interleaving and convolution code, in order to provide a quasi error free quality target at the input of the MPEG-2 demultiplexer. The convolution code r_2 is flexible, allowing the optimization of the system performance for a given MVDS transmitter bandwidth, e.g. for 33 MHz, for which the useful bit rate $R_u = r_1 \times r_2 \times R_s$ in Mbit/s is shown in Table 1. To achieve a very high power efficiency without excessively penalizing the spectrum efficiency, MVDS uses base band shaping by applying a square-root raised cosine filter with a roll-off factor of 0.35 ($B_{RF,3\,\mathrm{dB}}/R_S = 1.27$) and QPSK modulation with a symbol rate R_s of 26 Mbaud for 33 MHz. Table 1 exhibits the corresponding required signal-to-noise power ratio at the receiver input as a function of the convolution code r_2 [3].

Table 1
Useful bite rate R_u, energy per bit to noise power, E_b/N_0, and the required signal-to-noise power ratio $S/N_{req} = 2 \cdot r_1 \cdot r_2 \cdot E_b/N_0$ at the receiver input as a function of r_2 for QPSK and a 3-dB-bandwidth of 33 MHz

Parameter	$r_2 = 1/2$	$r_2 = 2/3$	$r_2 = 3/4$	$r_2 = 5/6$	$r_2 = 7/8$
R_u [Mbit/s]	24.0	31.9	35.9	39.9	41.9
E_b/N_0 [dB]	4.5	5.0	5.5	6.0	6.4
S/N_{req} [dB]	4.1	5.9	6.9	7.9	8.5

The allocated 2-GHz-frequency spectrum for digital MVDS is divided 1) into a downstream, including 96 RF-channels with a bandwidth of 33 MHz and a channel spacing of 39 MHz, and 2) into an upstream for return paths, having a bandwidth of 50 MHz at both ends of the 2 GHz band, respectively (Fig. 2). To double the capacity in cellular networks, use is made of vertical and horizontal polarization. Hence, the available spectrum might be separated into four groups of 24 RF-channels for cellular MVDS networks. The frequency error of each carrier shall not exceed ± 0.5 MHz at 40 GHz and the carrier output power shall not exceed 0.5 W (UK) or 1 W (GER) per channel. The power of spurious emission is given in [7]. Many aspects and issues of a one-way MVDS were already been discussed in the DIMMP-Project [2] and some proposals for the upstream issues are made by DAVIC [1] and DVB [4].



Group	*Downstream*	*Upstream*
1	40.55÷41.5 V	42.45÷42.5 H
2	40.55÷41.5 H	42.45÷42.5 V
3	41.5÷42.45 V	40.5÷40.55 H
4	41.5÷42.45 H	40.5÷40.55 V

Fig. 2. Channel plan for digital MVDS.

Typical transmitters in lower frequency bands uses IF-modulation, followed by a combiner, a solid state amplifier and transmit antenna. With the today's state of the art amplifiers, a maximum output power of 3 to 5 W is achievable at 40 GHz. However, typical values are 0.5 to 1 W (national limitations), providing an output power per channel of 10 to 20 mW for 24 RF-channels (low-power MVDS), taking into account a output back off of 5 to 6 dB. To increase the output power for achieving larger service distances, each RF-channel can be build up separately without a combiner (high-power MVDS). Figure 3 shows

a proposal from Philips broadband systems. After the multiplexer, each RF-unit consists of a QPSK-modulator, upconverter, amplifier and sector-horn antenna with a gain of approximately 15 dBi. A transmitter unit including 8+1 redundant of such RF-units with one local oscillator is shown in Fig. 3, too. Hence, 4 of such transmitter units are necessary for each MVDS cell.

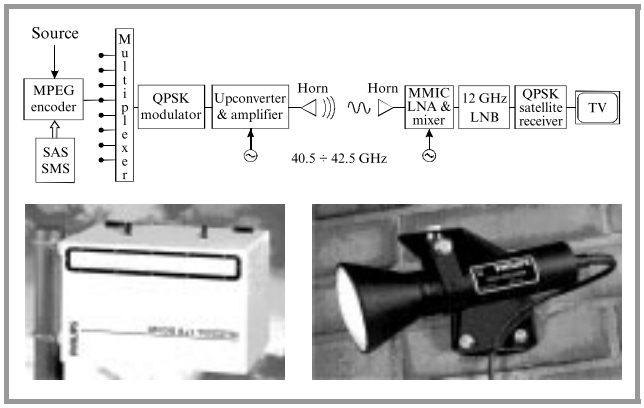


Fig. 3. High-power MVDS (each RF-channel is build up separately) from Philips broadband systems.

Close to the base station, a simple horn antenna can also be sufficient at the subscriber site, whereas pencil-beam antennas with very high gain are needed with increasing distance from the base station. This can be achieved by using paraboloidal or lens-horn antennas (Fig. 3), having a gain of around 30 to 36 dBi and 3-dB beamwidths of 5° to 2.5° for diameters of 10 to 20 cm.
Especially, the compact lens-horn antenna can easily be mounted on walls, is more aesthetic and causes much less mechanical problems due to wind forces than typical DBS antennas in the Ku-band. The receiving antenna is followed by a MMIC low noise amplifier and mixer, having a total noise figure of less than 6 dB. The mixer converts the received 40 GHz signal down to the 12 GHz region, where use can be made of standard DBS components (12-GHz LNB, a digital receiver and the TV).

3. LOS-propagation and service distances

The basic radio propagation characteristics and service area coverage of MVDS or LMDS systems is usually being limited to clear line of sight between the base station antenna and the customer antenna. Without atmospheric losses, the system gain is

$$G_s = P_{Tx} - L_{FTx} + G_T - L_R + G_R - L_{FRx} - N \quad [\text{dB}], \quad (1)$$

where

$$N = 10 \cdot \lg \left[k \cdot B_{RF} \cdot \left\langle \frac{T_A}{L_{FRx}} + T_F \left(1 - \frac{1}{L_{FRx}} \right) + \right.\right.$$
$$\left.\left. + (F-1) \cdot T_0 \right\rangle \right] \approx 10 \cdot \lg \left[k \cdot B_{RF} \cdot T_0 \right] + F \quad [\text{dB}] \quad (2)$$

is the thermal noise power, P_{Tx} the transmitter power, G_T and G_R the transmit and receive antenna gain, L_{FTx} and L_{FRx} the feeder losses, L_R the losses due to imperfect alignment, k the Boltzmann constant, B_{RF} the RF-bandwidth, F the noise figure of the receiver, T_A, T_F and $T_0 = 290$ K the antenna, feeder and the ambient temperature. The signal- or carrier-to-noise ratio C/N in the link budget is calculated by summing the system gain and the total path loss, including free-space attenuation $L_{fs} = 20 \cdot \lg \left[4\pi \cdot L/\lambda \right]$ with L for the distance and λ for the wavelength, clear sky attenuation A_{cs}, attenuation due to fog A_{fog} and rain attenuation $A_{h,v}(p)$:

$$C/N_{h,v}(p) = G_s - L_{fs} - A_{cs} - A_{fog} - A_{h,v}(p) \quad [\text{dB}]. \quad (3)$$

According to ITU-R Rep. 719-3 and 721-3, the specific attenuation during clear sky or due to fog are usually less than 0.25 dB/km or in between 0.05 to 1 dB/km for an average up to a high fog density at 42 GHz. The predominant factor affecting performance of LOS systems above 10 GHz is signal fading caused by heavy rain. Calculation starts from a knowledge of rain intensity statistics, e.g. those of ITU-R Rec. 837-1 for an average year, taking into account the climate zones $(CZ)E, H$ and K for variations across Germany, and an appropriate rain attenuation and cross-polarization model (Table 2) [5, 6].

Table 2
Rain intensity and specific attenuation (f = 42 GHz, horizontal polarization) for some percentages of time p of an average year and three climate zones E, H and K (ITU-R Rec. 837-1)

p [%]	Total time of a year	Point-rain rate R(p) [mm/h]			Specific attenuat. $A_h(p)$ [dB/km]		
		E	H	K	E	H	K
1	3.65 days	0.6	2	1.5	0.24	0.73	0.56
0.5	1.83 days	1.6	3.1	2.9	0.56	1.1	1.03
0.3	1.1 days	2.4	4	4.2	0.86	1.4	1.5
0.1	8.8 hours	6	10	12	2.0	3.2	3.8
0.03	2.6 hours	12	18	23	3.8	5.5	7.0
0.01	52 min.	22	32	42	6.7	9.4	12.1
0.003	16 min.	41	55	70	11.9	15.6	19.5
0.001	5 min.	70	83	100	19.5	22.8	27.0

In order to determine rain attenuation $A_{h,v}(p)$ use was made of a spatial rain model

$$A_{h,v}(p) = a_{h,v} \cdot \left[R(p) \right]^{b_{h,v}} \cdot L \ [\text{dB}]$$

for $R \leq 10$ mm/h and

$$A_{h,v}(p) = a_{h,v} \cdot \left[R(p) \right]^{b_{h,v}} \underbrace{\frac{1 - e^{-\gamma \cdot b_{h,v} \cdot \ln(R/10) \cdot L \cos\varepsilon}}{\gamma \cdot b_{h,v} \cdot \ln(R/10) \cdot \cos\varepsilon}}_{L \cdot \text{reduction factor } r} \quad [\text{dB}]$$

for $R > 10$ mm/h,

where $R(p)$ is the point rainfall rate (integration time of 1 min), p the exceeding probability of an average year, L the path length, $a_{h,v}$; $b_{h,v}$ the regression coefficients, depending on polarization, temperature, frequency and the drop size distribution, h and v the symmetry axes of the raindrop, ε the elevation angle and $\gamma = 1/14$ a fitting parameter for the rain profile (spatial decrease). Since the Laws-Parsons drop-size distribution characterizes the median values of statistical measurements well, the corresponding regression coefficients recommended by ITU-R Rep. 721-3 were used to determine cumulative distributions of an average year at 42 GHz. Because of paper limitations, all figures below are restricted to calculations for *CZ H* and horizontal polarization (worst case). For others see [5].

In order to illustrate trends in service areas, typical system parameters were assumed in the link budget of a digital MVDS: required $C/N = 6.8$ dB, $P_{Tx} = 0.5$ W, $G_T = 15$ dBi for a 64°-sector horn, $G_R = 32$ dBi for a lens-horn antenna, $F = 6$ dB, $B_{RF} = 33$ MHz, $L_R = 0.5$ dB, $L_{FTx} = 1$ dB and $L_{FRx} = 0.5$ dB. Figure 4 illustrates the coverage area of a 64°-sector horn transmit antenna of an elliptical, nearly circular shape. It is placed on the periphery has an azimuth pattern shown in Fig. 4. The effect of rain is quite apparent. For low rain rates, which correspond to low system availability percentages $(1 - p)$ in %, the service area is large. Conversely, the higher rain rate, or percentage availability, the smaller service area. To ensure that the target performance is maintained, e.g., for 99.9% of the time, the maximum service distance is nearly 6 km, compared to 27 km without rain. Because of this difference, an adaptive power control strategy seems to be reasonable, to reduce interference and EMC problems. The maximum service distance or the system margin at a given distance can be depicted from Fig. 5, where C/N is plotted vs. the service distance. For an average required C/N of 7 dB, the maximum service distance is approximately 6 or 3 km for an availability of 99.9 or 99.99% of the time.

Figure 6 exhibits the maximum service distance versus transmitter power per RF channel for a required C/N of 6.8 dB. A low-power MVDS with 10 to 20 mW meets the 99.9% criterion at a maximum LOS range of 2.5 to 3 km and a high-power MVDS with 0.5 to 1 W at 5.5 to 6 km.

MVDS especially in industrialized countries will only be competitive to CATV and DBS, if it can provide service-on-demand (SoD), i.e. if it can become a low-cost integrated medium with high-speed access to subscribers for interactive multimedia applications. For these interactive services an upstream or return path is needed from subscribers to the base station. Hence, interactive LMDS includes 1) a unidirectional forward broadcast path (FBP) for video, audio and data distribution, 2) a bi-directional interaction system composed of a forward interaction path (FIP) and a return interaction path (RIP) [1, 4]. The required data rates for interactive services are very dependent on the kind of services being requested. For example, video on demand, tele-shopping, or remote learning needs high bit rates of 2 to 6 Mbit/s in the FIP, whereas 20 kbit/s are sufficient in the RIP. In contrast, a video conference or an exchange of multimedia data needs in both interaction paths several Mbit/s. There are different strategies to meet these different data rate requirements and future service requests of customers 1) RIP and/or FIP are implemented out of band, offering low data rates via an external medium and, 2) RIP and/or FIP are embedded in the 2-GHz MVDS spectrum (in-band), able to provide higher bite rates up to several Mbit/s. The first option causes problems in combining the

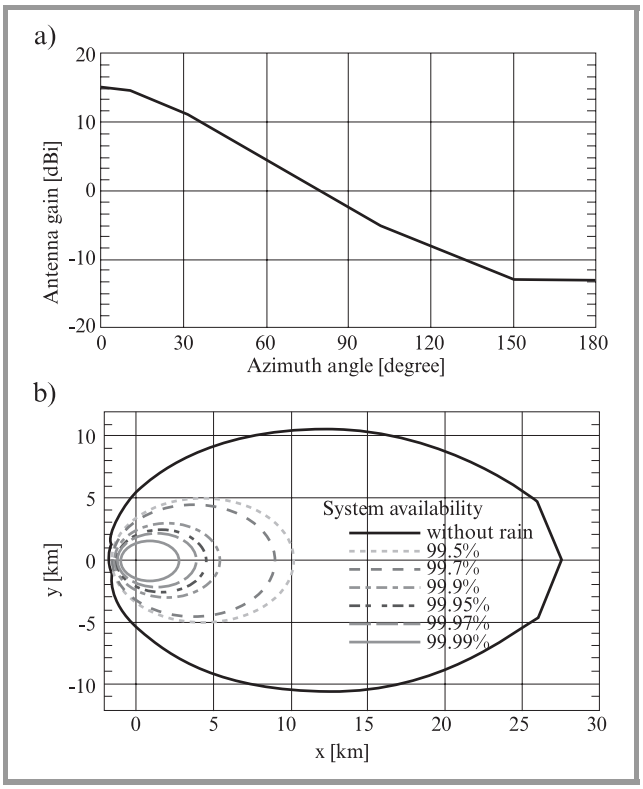


Fig. 4. (a) Azimuth pattern of a 64°-sector horn transmit antenna and (b) coverage area at 42 GHz for *CZ H* and horizontal polarization.

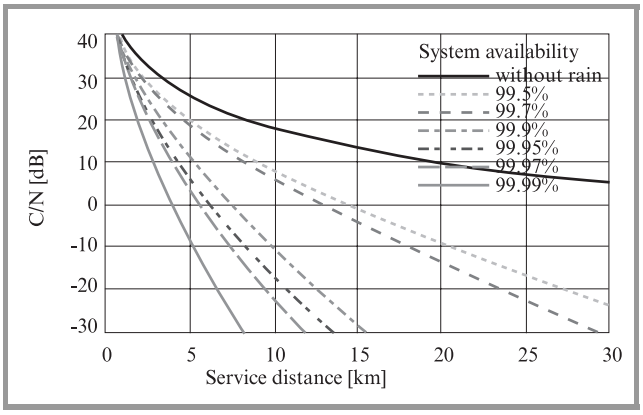


Fig. 5. C/N versus service distance for *CZ H* and horizontal polarization.

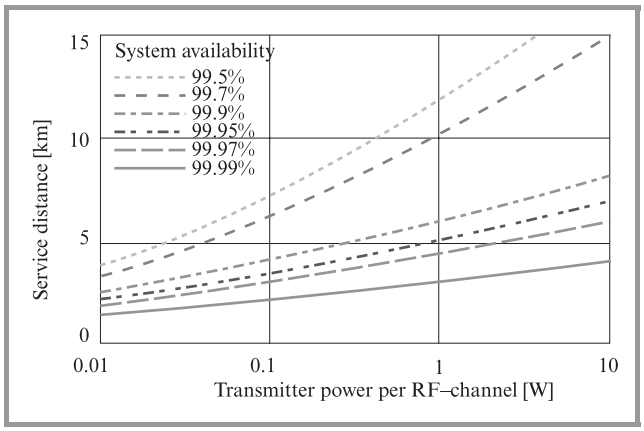


Fig. 6. Service distance versus the transmitter output power per RF channel for $C/N_{req} = 6.8$ dB, *CZ H* and horizontal polarization.

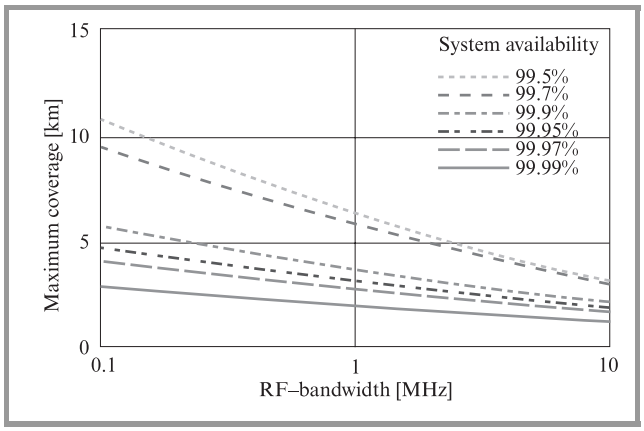


Fig. 7. Maximum coverage between customer station and base station versus RF-bandwidth of the RIP for an output power of 10 mW.

connections of FBP with external FIP and/or RIP, whereas transporting inter-active broadband services over a single wireless link seems to be more attractive from a customers point of view (single connection) and, it can be the only or at least a cost-effective solution for the competitors of the established operators in some areas. Therefore, solely the last option will be considered in this paper.

In order to keep the expenses of the customer station as low as possible, its complexity and transmitter output power should be low. Hence, robust modulation techniques such as QPSK or DQPSK without the need of a FEC is a good candidate. Figure 7 shows the maximum distance or coverage between customer stations and base station as a function of the RF-bandwidth of the RIP for an output power of 10 mW, where the 99.9% availability criterion meets the maximum LOS distance at 3.5 and 2.5 km for a bandwidth of 1 and 8 MHz. The distance can be increased up to 4 and 6 km for P_{Tx} of 100 mW [5].

4. The single-cell LMDS planing tool

The single-cell coverage prediction takes place in three steps as shown in Fig. 8. First, the number of visible households is calculated in order to select the best site for the base station, where usually only few sites are at one's disposal. Therefore, the visible parts of all buildings within the considered area have to be evaluated serving as input information for the household-estimation model, which will be presented below. All these calculations are based on 3D databases of both, buildings and vegetation. By the way, vegetation is here considered as blocking obstacles. This seems to be reasonable, because of large transmission losses of more than 30 dB above 26 GHz and the large time variations in the received power of up to 10 dB due to wind and seasons. Thus, a coverage through vegetation is not practicable at millimeterwaves.

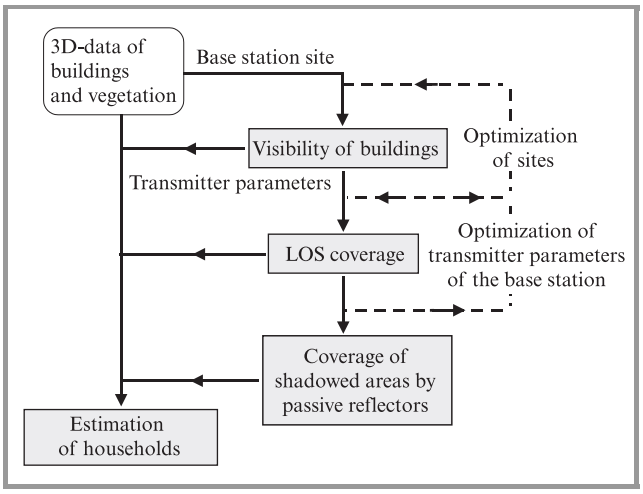


Fig. 8. Scheme of the single-cell LMDS planing tool.

Having selected the best site, the parameters of the base station's transmitter can be optimized in a second step, particularly the transmitter power and the orientation of the antenna in both, azimuth and elevation, by calculating the LOS area coverage under rainy conditions for certain rain climate zones. In a final step, specific reflectors can be used to increase the number of households by covering shadowed areas near the base station [9].

4.1. The household-estimation model

Firstly, it has to be defined what a potential customer might be. This is quite difficult, because this may vary with the provided application or service. Therefore, in our approach, a household is considered as an individual apartment or accommodation unit. This assumption is valid for most of the private buildings in urban or suburban areas, but will not be correct for public and business buildings representing just one customer, independent of the building size. To avoid any mistakes, additional residential information are needed.

The number of households within a building can be assessed either by the space-volume or the external surface of a house. Because a household is counted to be covered if at least a part of the corresponding surface of the building is covered, it seems to be reasonable to choose a surface-oriented algorithm rather than a space-volume-oriented method to estimate the number of households. Therefore, a building is represented by its surface consisting of walls and roof plains as shown in Fig. 9.

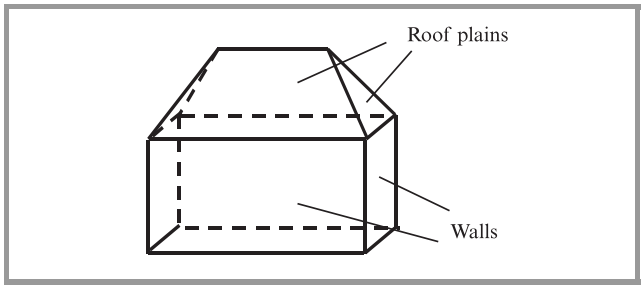


Fig. 9. Surface-oriented model of a building.

To estimate the number of covered households $N_{HH\,\mathrm{cov}}$, the covered parts of the building surface is needed. This input information has to be delivered by the planing tool. Hence, it has to be taken into account that the number of covered households depends not only on the size of the covered area but also strongly on the location of the covered parts. An example is shown in Fig. 10.

Therefore, the basic idea behind our household-estimation model is to choose the most probable location of a household in a building. Firstly, this is done by dividing a building into floors using an average floor height h_{floor}. In addition, the roof is treated as a floor, however with a weighting factor less than 1 or even neglected, depending on the declination of the roof. Furthermore, every wall is subdivided into several sections according to an average estimated length of a household l_{HH} as shown in Fig. 11. Of course several sections can belong to just one household.

Thus, by considering the floors separately, the calculation of the number of households N_{HH} and the number of covered households $N_{HH\,\mathrm{o}}$ becomes to

$$N_{HH} = \sum_{i}^{N_{\mathrm{floor}}} N_{HH,\mathrm{floor}}(i) \tag{4}$$

$$N_{HH\,\mathrm{cov}} = \sum_{i}^{N_{\mathrm{floor}}} N_{HH\,\mathrm{cov},\mathrm{floor}}(i)\ . \tag{5}$$

For the calculation of the number of covered households of a certain floor $N_{HH\,\mathrm{cov},\mathrm{floor}}$ it has to be taken into account, that one household can have several covered sections as shown in Fig. 12, where the household has to be considered only once.

A deterministic approach to avoid a multiple consideration of one household would be to define households and

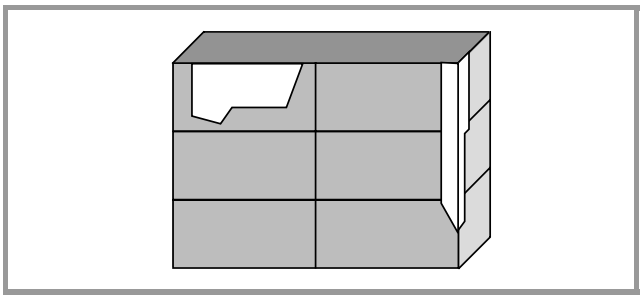


Fig. 10. Building with two visible parts of nearly the same size, but covering different numbers of households.

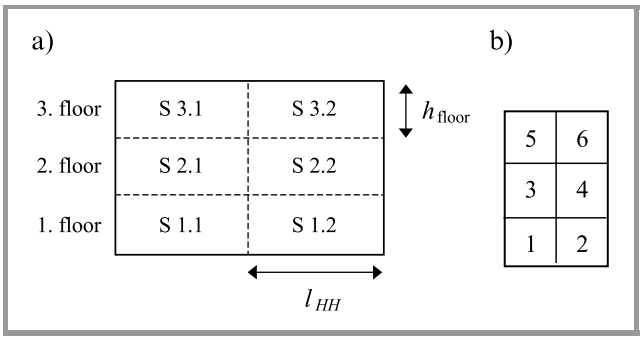


Fig. 11. Front sight of a single building wall (a) divided into floors and sections as well as the corresponding ground plane of the building with 6 households per floor (b).

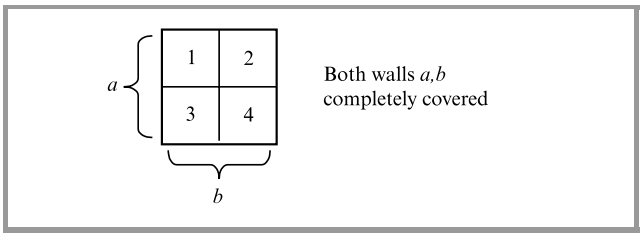


Fig. 12. One household (no. 3) with two covered sections.

assign the boundary sections to them. Then, also the number of households could easily be calculated. This method would be based on the assumption that the real households are essentially represented by sections, which generally is not valid. Therefore, in the following a more statistical algorithm is suggested for the estimation of $N_{HH,\mathrm{floor}}$ and $N_{HH\,\mathrm{o},\mathrm{floor}}$, respectively.

Obviously, there must be a correlation between the number of existent households per floor $N_{HH,\mathrm{floor}}$ and the number of sections of a floor $N_{S,\mathrm{floor}}$, depending on the shape of the building. For this correlation a good approximation is given by

$$N_{HH,\mathrm{floor}} = N_{S,\mathrm{floor}} - N_{\mathrm{convex}}\ , \tag{6}$$

where N_{convex} is the number of convex corners of a building.

Only the covered walls of the building are considered for the estimation of the covered households of a floor $N_{HH\,\mathrm{o},\mathrm{floor}}$.

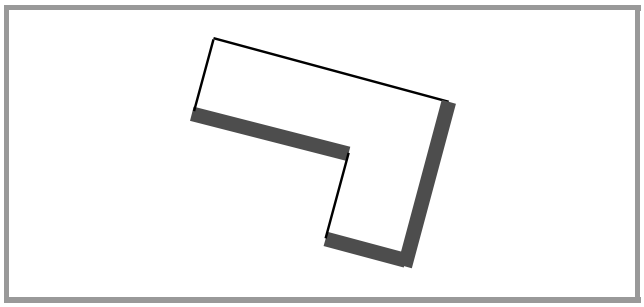


Fig. 13. Building with three covered walls (thick lines) and one covered convex corner $(N_{\text{convex,cov}} = 1)$.

A wall is counted to be covered, if only a part of it is covered. The percentage of covered sections of these walls are calculated and multiplied with the total number of all those households $N_{HH\,\text{walls,o}}$ with covered walls. Thus, the number of covered households of a floor is approximated by

$$N_{HH\,\text{cov,floor}} \approx \frac{N_{S\,\text{cov,floor}}}{N_{S\,\text{walls,cov}}} \cdot N_{HH\,\text{walls,cov}} \,. \tag{7}$$

Similar to the estimation of the existent households, the number of households $N_{HH\,\text{walls,cov}}$ with covered walls is computed from the number of sections $N_{S\,\text{walls,cov}}$ and the number of convex corners $N_{\text{convex,cov}}$ of the covered walls:

$$N_{HH\,\text{walls,cov}} = N_{S\,\text{walls,cov}} - N_{\text{convex,cov}} \,. \tag{8}$$

Here, a convex corner has to be formed by two adjacent covered walls as shown in Fig. 13.

The quality of our results of the presented model depends very much on the choice of the fitting parameters h_{floor} and l_{HH}, which have to be assessed on average from corresponding residential data. Therefore, a real validation of the suggested models requires a very large number of data, which has not been available yet. However, tests with buildings of the most common shapes show that the estimation of the number of existent households works quite well. In general, the results are getting more and more incorrect, the more complex the structure of a building is. However, because the estimation model for the number of covered households of a building and the model for the number of existent households of a building are quite similar, the quotient of both numbers, representing the percentage of coverage, will be quite accurate, even if the single estimations are not precise.

4.2. Design of passive reflectors to cover shadowed areas

In order to increase the number of covered households, particularly near the base station, the possibility to cover shadowed areas by using shaped metallic reflectors has been investigated for 42 GHz. In [11], use was already made of plane reflectors acting as passive repeaters. But covering additional households needs usually a much wider beamwidth, thus a reflector has to be convex. In fact, the

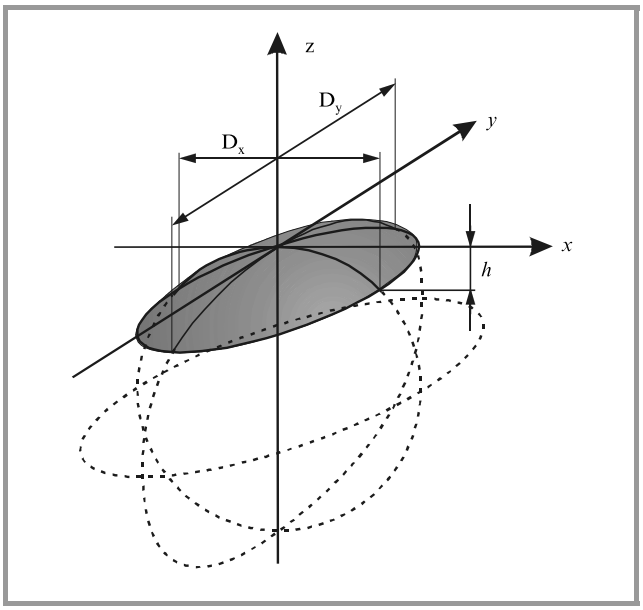


Fig. 14. Shape of the passive reflector.

radiation pattern of a passive reflector has to be adapted to a local situation, in order to cover as many customers as possible or aimed customers to be covered with LMDS, e.g. a shadowed street or building block. The required beamwidth in azimuth and elevation needs usually only a weak curvature of the reflector, i.e. only a slight deviation from a plane reflector. Hence, it is sufficient to consider merely a part of an ellipsoid with the diameters D_x and D_y or a part of a sphere $(D_x = D_y = D)$, both with the height h of the cap as shown in Fig. 14.

For both shapes, a pattern analysis using PO has been carried out in dependence of the height h, the diameters D_x and D_y as well as the angle of incidence of the incoming wave from the base station. In order to accelerate the computer speed to analyze the impact of passive reflectors applied in LMDS cells as well as for practical reasons, simple rules have been derived from the PO analysis for the design of metallic reflectors with a desired radiation pattern and gain [9].

The radiation patterns of a spherical reflector for different curvatures are shown in Fig. 15.

As you can see from Fig. 15, there are ripples in the main beam so that there is no real half power beamwidth (HPBW). Nevertheless, you can define a kind of average beamwidth Θ_{BW}, which seems to be proportional to the height h of the reflector, whereas the power density p decreases with h^{-2}. Including the influence of the diameters D_x and D_y you can find the approximations:

$$\Theta_{BW,\,\text{Az}} \approx 890^\circ \cdot \frac{h}{D_x} \;[^\circ] \tag{9}$$

$$\Theta_{BW,\,El} \approx 890^\circ \cdot \frac{h}{D_y} \;[^\circ] \tag{10}$$

for 3 mm $< h <$ 15 mm and 0.3 m $< D_x, D_y <$ 1 m.

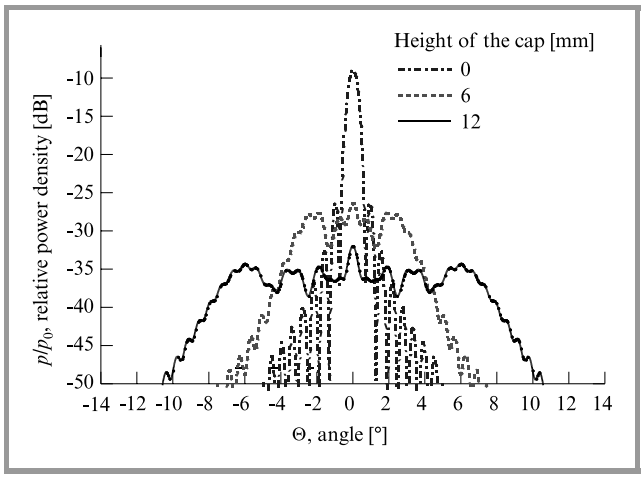


Fig. 15. Radiation pattern in a distance r = 150 m of a spherical reflector with infinite conductivity and a diameter of $D_x = D_y =$ 0.7 m for different heights h of the spherical cap and perpendicular incidence.

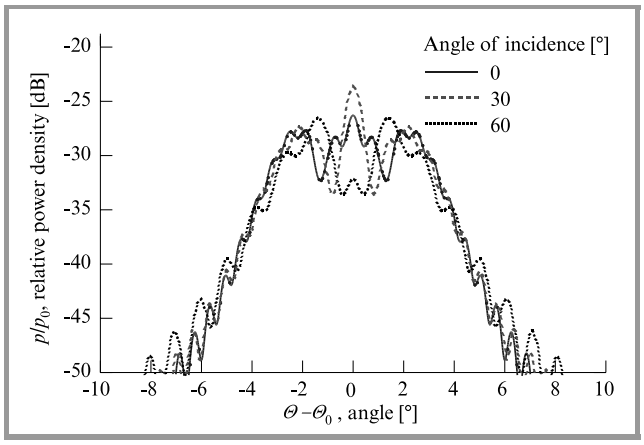


Fig. 16. Radiation pattern in a distance r = 150 m of a spherical reflector ($D_x = D_y = 0.7$ m, $h = 0.006$ m) with infinite conductivity and for different angles of incidence Θ_0.

Thus, the height h of a ellipsoidal reflector is determined by the maximum diameter corresponding to the smaller beamwidth. Another important result of the PO simulations is: varying the angle of incidence Θ_0 in a range of up to $\Theta_0 = 60°$ has nearly no impact on the average beamwidth Θ_{BW}, it causes only slightly different ripples as shown in Fig. 16.

The power density $p(r)$ at the observation point can easily be calculated by the following far-field expression

$$p(r) = P_R \cdot G \cdot \frac{1}{4\pi \cdot r^2}\,, \tag{11}$$

using an approximation for the "gain" of the reflector

$$G \approx \frac{4\pi}{\Theta_{BW,AZ} \cdot \Theta_{BW,EL}} \tag{12}$$

and the reflected power

$$P_R \approx p_0 \cdot \frac{D_x \cdot D_y \cdot \pi}{4}\,, \tag{13}$$

where p_0 denotes the power density from the source at the reflector surface and r the distance from the reflector.

This approximation has to be refined especially near the reflector. After introducing an additional empirical factor and defining a breakpoint r_1, we actually found the following approximations from the PO analysis:

$$\frac{p(r)}{p_0} = 1 \qquad \text{for } r < \frac{D_x \cdot D_y}{1.28\,\text{m}} \tag{14}$$

$$\frac{p(r)}{p_0} = \frac{1}{1.28\,\text{m}} \cdot \frac{D_x \cdot D_y}{r} \sim \frac{1}{r} \qquad \text{for } r < r_1 \tag{15}$$

$$\frac{p(r)}{p_0} = \frac{1.34}{\text{m}} \cdot \frac{1}{h} \cdot \left(\frac{D_x \cdot D_y}{r}\right)^2 \sim \frac{1}{r^2} \qquad \text{for } r > r_1 \tag{16}$$

with the breakpoint at

$$r_1 = 1.71 \cdot \frac{D_x \cdot D_y}{h}\,. \tag{17}$$

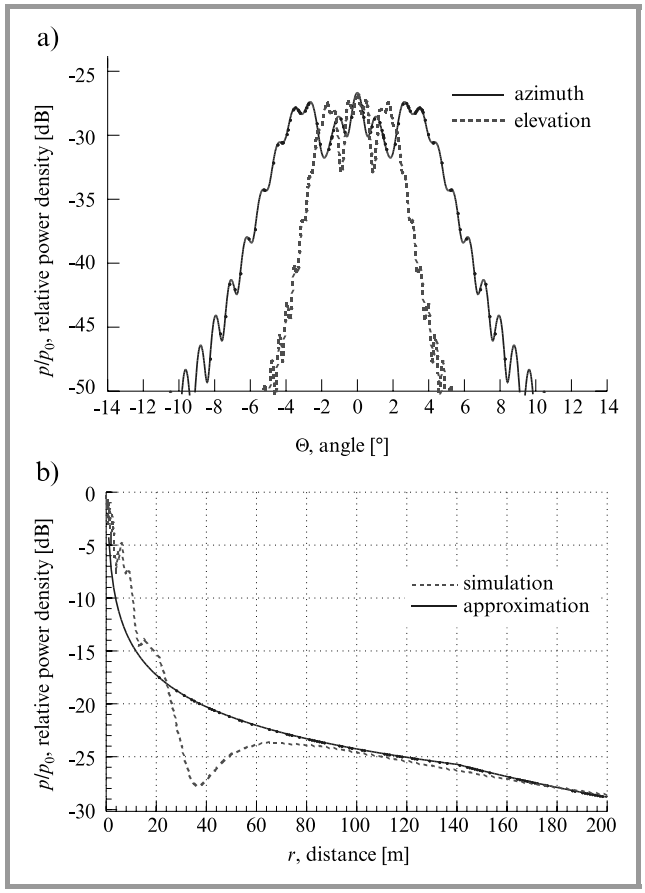


Fig. 17. (a) Radiation pattern in a distance r = 150 m and (b) relative power density as a function of the distance for an ellipsoidal reflector with D_x = 0.5 m and D_y = 1 m and height h = 0.006 m.

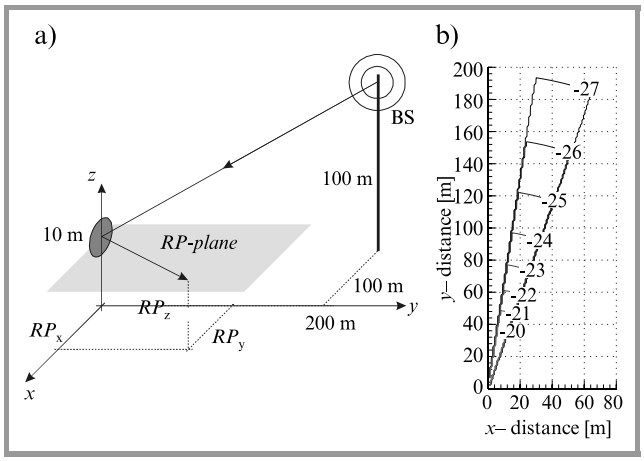


Fig. 18. Example of covering a shadowed street with a passive reflector: (a) layout of the arrangement: base station, metallic reflector ($D_x = 0.5$ m, $D_y = 1.0$ m, $h = 0.006$ m) and the receiving-point plane (RP), which is located 10 m above ground; (b) relative power density profile p/p_0 in dB in azimuth.

For a typical ellipsoidal reflector, Fig. 17 shows a good agreement of these approximations with the result of the PO simulations, the radiation pattern with PO (Fig. 17a) and the relative power density as function of the distance with PO (Fig. 17b). Obviously, there is a deep fading near the reflector. However, the fading only occurs very close to the boresight axis ($\Theta < 1°$) and can therefore usually be neglected.

In practice the size of the reflector is for aesthetic reasons limited to about 1 m. Because of the limited gain, the range of using passive reflectors is restricted to a distance of about 1 to 2 km around the base station. The range of the reflector's covering itself along the boresight is then up to about 200 m (Fig. 18), depending on the reflector's curvature or required beamwidth and the system margin [9, 10]. Thus, the usage of passive reflectors to cover shadowed areas is in principal possible and can be a cheap solution in certain areas, e.g. a shadowed street or building blocks.

5. Conclusions

A software tool for single-cell LMDS coverage prediction has been implemented based on 3D databases of buildings and vegetation. In order to provide the operator with information about the profitability of a cell, the desired number of potential customers has been approached by using simple household-estimation models. Apart from mere LOS considerations, also the coverage via shaped metallic reflectors has been investigated. Simple designing rules for these reflectors have been derived from detailed simulations at 42 GHz using PO. It shows that a passive reflector can be a cheap solution to cover shadowed areas for distances of about 1 to 2 km around the base station. As a next step, the multipath propagation effects caused by reflections and scattering on buildings and vegetation as well as co-channel interference have to be considered in the software planing tool. This has already been started in [12] and will be investigated in more detail in the near future using ray tracing techniques.

References

[1] H. Brüggemann, "DAVIC – Neue, weltweite Spezifikationen für digitale audiovisuelle und multimediale Dienste", *Telekom Praxis*, Nr. 2/96, 1996.

[2] "RACE DIMMP Workshop Papers", Metz, 1995.

[3] EN 300 748, "Digital Video Broadcasting (DVB); Multipoint Video Distribution Systems (MVDS) at 10 GHz and above", ETSI, 1997.

[4] EN 301 199, "Digital Video Broadcasting (DVB); Interaction Channel for Local Multipoint Distribution Systems (LMDS)", ETSI, 1998.

[5] R. Jakoby and M. Grigat, "MMDS zur Erweiterung von BK-Netzen", Research Report from TZ of Deutsche Telekom AG, 1996.

[6] R. Jakoby and F. Rücker, "Analysis of depolarisation events", PIERS, ESA Center, Noordwijk, CD-ROM, 1994.

[7] United Kingdom Radiocommunication Agency, "Performance Specification for Analogue/Digital Multipoint Video Distribution Systems (MVDS) Transmitters and Transmit Antennas Operating in the Frequency Band 40.5 – 42.5 GHz", United Kingdom RA, MPT 1550/1560, 1993/1995.

[8] U. Reimers, *Digitale Fernsehtechnik; Datenkompression und Übertragung für DVB*. Berlin-Heidelberg: Springer Verlag, 1995.

[9] R. Jakoby, G. Bauer, A. Hayn, and J. Freese, "Entwicklung von Ausbreitungsmodellen und Softwaremodulen zur Abschätzung der MWS-Versorgung unter Berücksichtigung von abgeschatteten Bereichen", Research Report for T-Nova, Deutsche Telekom, Germany, 1999.

[10] R. Jakoby and A. Hayn, in *Proc. 8th URSI Symp.*, Aveiro, Portugal, 1998.

[11] A. Vigants, "Space Diversity on a 6-GHz Path with Billboard Reflectors", ICC, 12B.1.-4, 1974.

[12] A. Hayn, R. Bose, and R. Jakoby, "Multipath propagation and LOS interference studies for LMDS architecture", in *Proc. ICAP 2001*, Manchester, UK, 2001.

Georg Bauer received his Dipl.-Ing. (M.Sc.) degree from the University of Stuttgart, Germany, in 1998. He is currently working as a Research Assistant pursuing a Ph.D. degree at the Institute of Microwave Engineering of the Technical University of Darmstadt, Germany. His research interests include wave propagation and broadband wireless access.

e-mail: g_bauer@hf.tu-darmstadt.de
Technische Universität Darmstadt
Institut für Hochfrequenztechnik
Fachgebiet Mikrowellentechnik
Merckstrasse 25
D-64283 Darmstadt, Germany

Jens Freese received the Dipl.-Ing. (M.Sc.) degree from the University of Siegen, Germany, in 1999. Since October 1999, he is working as a Research Assistant towards the Dr.-Ing. (Ph.D.) degree at the Institute of Microwave Engineering of the Technical University of Darmstadt, Germany. His research interests include adaptive antennas and microstrip antennas for millimeter wave applications.
e-mail: freese@hf.tu-darmstadt.de
Technische Universität Darmstadt
Hochfrequenztechnik
Fachgebiet Mikrowellentechnik
Merckstrasse 25
D-64283 Darmstadt, Germany

Rolf Jakoby – for biography, see this issue, p. 19.

Paper

Fuzzy logic classifier for radio signals recognition

Jerzy Łopatka and Maciej Pędzisz

Abstract — This paper presents a new digital modulation recognition algorithm for classifying baseband signals in the presence of additive white Gaussian noise. Elaborated classification technique uses various statistical moments of the signal amplitude, phase and frequency applied to the fuzzy classifier. Classification results are given and it is found that the technique performs well at low SNR. The benefits of this technique are that it is simple to implement, has generalization property and requires no apriori knowledge of the SNR, carrier phase or baud rate of the signal for classification.

Keywords — modulation recognition, fuzzy logic.

1. Introduction

The problem of automatic classification of the applied modulation type of a digital transmission has received international scientific attention for over a decade now.

An automatic modulation classifier is a system that automatically identifies the modulation type of the received signal. It is an intermediate step between signal interception and information recovery. When the modulation scheme of a received signal is identified, an appropriate demodulator can be selected to demodulate the signal and then recover the information.

Modulation type classifiers play an important role in some communication applications such as signal confirmation, interference identification, surveillance, monitoring, spectrum management, electronic warfare, military threat analysis, or electronic counter-counter measure [6].

In this paper, we propose a new pattern recognition approach based on various statistical characteristics of the signal amplitude, phase, and frequency. Theoretically, feature patterns should produce independent and isolated classes. In practice, the classes are overlapped and it is hard to classify observed pattern to only one class. In such a case, there is a need to specify the grade of membership to each class. This method is naturally implemented as a fuzzy classifier.

Extracted features are applied to the Mamdani fuzzy classifier [4]. Chosen features create a low order modulation type model. This increases the generalization capability of the model, reduce computational complexity, simplify fuzzy rules and improve decision process.

In Section 2, we give signal model and problem formulation. In Section 3, the feature extraction process is shown and in Section 4 the classification algorithm is discussed. Section 5 illustrates the experimental results and conclusions are presented in Section 6.

2. Signal models

Let us assume knowledge of the carrier frequency. Received baseband signal can be expressed as

$$s(t) = x(t)e^{j\Theta_c} + n(t)\,, \tag{1}$$

where Θ_c is the carrier phase and $n(t)$ is a complex white Gaussian noise.

For quadrature amplitude modulation (QAM) signals,

$$x_{QAM}(t) = \sum_{i=1}^{N}(A_i + jB_i)u(t - iT), \tag{2}$$

where $A_i, B_i \in \{2m - 1 - M,\ m = 1, 2, \ldots, M\}$. When B_i in Eq. (2) is zero, then QAM signal becomes amplitude shift keying (ASK) signal.

For phase shift keying (PSK) signals,

$$x_{PSK}(t) = \sqrt{S}\sum_{i=1}^{N} e^{j\varphi_i}u(t - iT), \tag{3}$$

where $\varphi_i \in \{\frac{2\pi}{M}(m-1),\ m = 1, 2, \ldots, M\}$.

For frequency shift keying (FSK) signals,

$$x_{FSK}(t) = \sqrt{S}\sum_{i=1}^{N} e^{j(\omega_i t + \Theta_i)}u(t - iT), \tag{4}$$

where $\omega_i \in \{\omega_1, \omega_2, \ldots, \omega_M\}, \Theta_i \in (0, 2\pi)$.

In Eqs. (2), (3) and (4), S is the signal power, N is the number of observed symbols, T is the symbol duration and $u(t)$ is the pulse shape of duration T.

3. Feature extraction

On the basis of Eqs. $(2 \div 4)$, one can see that for specific modulation type, one or more parameters are being changed. In ASK case – amplitude is being modulated, for FSK signals – frequency, QAM is the case where both amplitude and phase are being changed. Number of parameters and the way in which they are changed are specific to each modulation type. They may be described by a set of statistical parameters related to its moments and distributions. After initial selections we have chosen three of them as most comprehensive way of signal description.

The first feature we propose to use is a kurtosis of a signal envelope. Kurtosis is a measure of how outlier-prone the distribution is. The kurtosis of the normal distribution is 3. Distributions that are more outlier-prone than the normal distribution have kurtosis greater than 3, distributions that are less outlier-prone have kurtosis less than 3. Theoretically, the envelope distribution of PSK and FSK signals is

a Rician distribution. In practice, this distribution may be approximated by Gaussian distribution for SNR > 10 dB. In that case, kurtosis for these signals will be approximately equal 3. If we receive ASK signal, envelope distribution will be a mixture of Rayleigh and Rician distribution, and kurtosis will be approximately 1. In M-ary ASK and QAM cases, distributions will be a combination of distributions given above.

Second feature we extract from the phase histogram of the analyzed signal. One well-known asymptotic expression of a phase probability density function (PDF) under assumption that $\sigma_w \to 0$ is a Tikhonov probability density function:

$$f_\phi(\phi_0) \cong \frac{exp\left[2\gamma\cos(\phi_0)\right]}{2\pi I_0(2\gamma)};\ -\pi < \phi_0 \le \pi;\ \gamma = \frac{A^2}{2\sigma_w^2}\,, \quad (5)$$

where σ_w is the noise standard deviation, $I_0[\cdot]$ is the zero-order modified Bessel function of the first kind. The approximation in (5) is considered good for values of $\gamma > 6$ dB and is considered fair for γ around 0 dB. For M-ary PSK signals, $M = 2^\alpha, \alpha = 0, 1, 2...$, the PDF of ϕ_α is a sum of the noncentral Tikhonov functions [6]:

$$f_\phi(y;\alpha) = \frac{1}{2^\alpha}\sum_{k=1}^{2^\alpha}\frac{exp\{[2\gamma\cos[y-\eta_k(\alpha)]\}}{2\pi I_0(2\gamma)}\,, \quad (6)$$

where $\eta_k(\alpha)$ is the phase of k-th phase state and can be expressed as

$$\begin{aligned}\eta_k(\alpha) &= \frac{(2k-2^\alpha-1)}{2^\alpha}\,,\\ k &= 1,2,\ldots,2^\alpha\,,\\ \alpha &= 0,\ldots,\log_2 M\,.\end{aligned}$$

The number of peaks in the phase PDF, indicates the number of signal phase states. Generally, $f_\phi(y;\alpha)$ approaches $1/2\pi$ for SNR $\to -\infty$ dB or $\alpha \to \infty$. When we compute FFT magnitude of the phase histogram, it is easily seen that for M-ary PSK and QAM signals there are spectral lines that indicates M-ary number. For ASK and FSK signals, there are no spectral lines and the power spectral density (PSD) is much smoother. If we compute derivative of that PSD, this effect will be more visible. Figure 1 shows both: PSD and its derivative for ASK and 4DPSK signals. If in the received signal, phase is a parameter being modulated (PSK, QAM), the variance of the PSD derivative should be large. In the other case (ASK, FSK), variance should be smaller.

Third feature we propose is a mean signal frequency. Suppose that the receiver is tuned to signal center frequency. Digitally modulated signals we can model as a set of complex exponentials. In this case, one can estimate signal PSD by Burg's method [3]. This method fits an autoregressive (AR) model to the signal by minimizing the forward and backward prediction errors while constraining the AR parameters to satisfy the Levinson-Durbin recursion. The PSD estimated in this way is much smoother than estimated by nonparametric methods (Welch, FFT), and the spectrum consists only the meaningful part of the signal. This meaningful part means surroundings of the carier frequencies

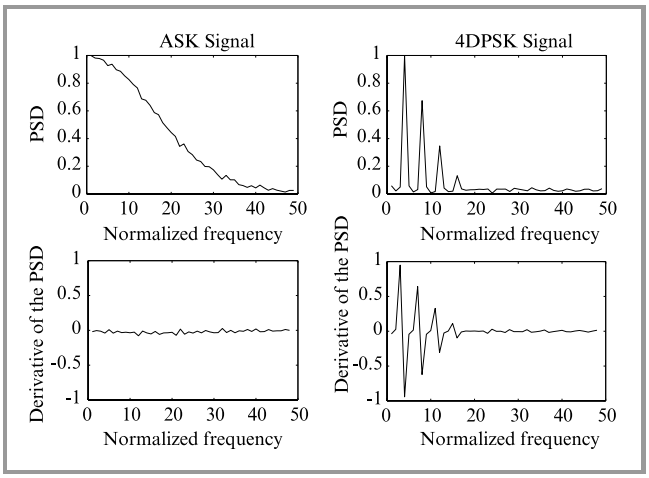


Fig. 1. PSD and its derivative (v = 200 Bd; SNR = 15 dB).

in ASK, PSK and QAM signals. If the assumption concerning proper tuning of the receiver is reached, then the mean of the absolute value of instantaneous frequency is approximately zero for ASK, PSK, and QAM signals, and is larger for FSK and MSK signals. Its value depends on the symbol rate and the frequency deviation.

These features create overlapped classes, and it is hard to say what means for example "large variance". To solve this problem, we conducted experiments to define the membership functions for each feature. There are three membership functions for the amplitude, two for the phase, and two for the frequency feature. We also defined membership functions for the classifier output. This output produces fuzzy decisions.

Applied modulation type may be in some degree ASK, PSK or FSK. Membership functions used for the inputs are sigmoidal:

$$f(x,a,c) = \frac{1}{1+e^{-a(x-c)}}\,, \quad (7)$$

where a is the stretch parameter, and c is the mean parameter.

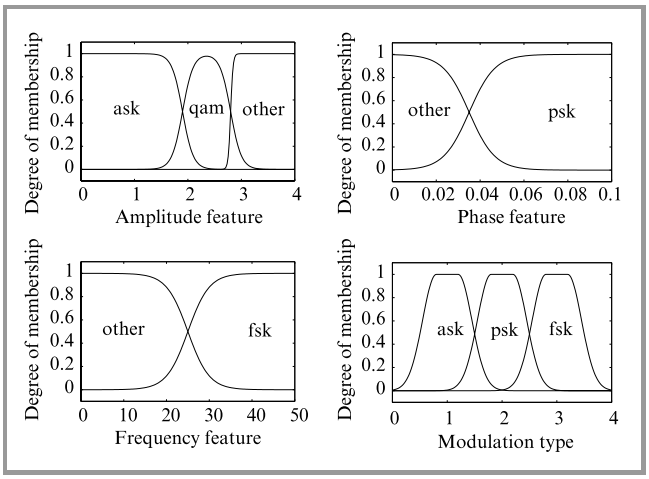


Fig. 2. Membership functions for the inputs and the output.

Membership function for the output is a generalized bell-shaped one:

$$f(x,a,b,c) = \frac{1}{1+|\frac{x-c}{a}|^{2b}}\,, \tag{8}$$

where c locates the center of the curve, a is the stretch parameter, and b is related to its width. Figure 2 shows all these membership functions for the inputs and the output.

4. Classification

After the features are extracted, the classification process can be applied. Because membership functions overlap each other, the fuzzy rules of classification and inference were used.
We constructed five rules to specify which modulation type has been applied in the received signal. We also used symbolic descriptions for the features and for the fuzzy sets. *A*, ϕ, ω are the features extracted from signal envelope, phase and frequency, whereas *ask, psk, qam, fsk, mfsk* and *other* are fuzzy sets associated with appropriate membership functions. The proposed rules are as follows:

If (A is ask) and (ϕ is other) and (ω is other) then (type is ask)
If (A is qam) and (ϕ is psk) and (ω is other) then (type is ask)
If (A is qam) and (ϕ is psk) and (ω is other) then (type is psk)
If (A is other) and (ϕ is psk) and (ω is other) then (type is psk)
If (A is other) and (ϕ is other) and (ω is fsk) then (type is mfsk)

Weighting coefficients for these rules are assumed: 1, 0.5, 0.5, 1 and 1 respectively. The classifier is based on Mamdani fuzzy classifier [4] with following parameters:

- AND method: PROD;
- IMPLICATION method: MIN;
- AGGREGATION method: PROBOR;
- DEFUZZYFICATION method: CENTROID.

The answer of the classifier can be given either as the single output or as three-output. Single output gives the number, which indicate what type of modulation has been applied. This value ranging from 0 to 4, where 1 corresponds to ASK signal, 2 to PSK signal, 3 to FSK signal and combined modulation types are indicated as an intermediate numbers (i.e. QAM may be represented as 1.5). The three-output answer gives three numbers, which specify the degree of membership of the three basic modulation types (i.e. QAM may be represented as: 0.5, 0.5, 0).

5. Experimental results

In this section, we demonstrate the performance of the suggested algorithm. Our goal is to discriminate between ASK, 4DPSK, 16QAM and FSK signals. We assume that we have 8000 samples of the complex signal, sampling frequency is 8 kHz, baud rate is 200 Bd and in FSK case – frequency shift is orthogonal. We use three-output classifier answer, and four signals for tests (ASK, 4DPSK, 16QAM and FSK). As we increase SNR, the performance of the classifier improves. For SNR $>$ 5 dB, the classifier output is very good. When SNR is less than 5 dB then classifier performance is getting worse. The worst results we get for the 16QAM case: when SNR $=$ 0 dB, then analysed signal is classified as a PSK signal. We have to notice that classifier was trained only for SNR $>$ 5 dB (for distinguishing membership functions), and its behavior for SNR $<$ 5 dB results from its generalization property. Table 1 shows simulation results of estimated correct classification probabilities versus SNR for four tests signals and 100 repetitions.

Table 1
Probability of correct recognition versus SNR

SNR [dB]	0	2	4	6	8	10
ASK	0.91	0.99	1.00	1.00	1.00	1.00
4DPSK	0.85	0.90	0.95	0.98	1.00	1.00
16QAM	0.00	0.31	0.73	0.95	0.99	1.00
FSK	0.99	1.00	1.00	1.00	1.00	1.00

6. Conclusions

In this paper a new method for the automatic modulation classification of ASK, PSK, QAM and FSK signals has been presented. This method requires no apriori knowledge of the SNR, carrier phase, or baud rate of the signal. Simulation results proved that the elaborated algorithm, using proposed set of the features is very robust with respect to SNR. The robustness of the classifier follows from its fuzzy structure. The cost of this achievement lies in the additional complexity of the inference process. Simulations showed that for SNR larger than 5 dB, classifier works properly. Soft decisions generated by the classifier carry two types of information: the applied modulation type and the degree of membership of this type. It can be used as a parameter in the next stage of the classification process, identification of the constellation shape and M-ary number in PSK and QAM signals or determining number of frequencies and shifts in FSK signals. It allows to create intelligent radio links, efficient monitoring and control systems.

References

[1] J. Aisbett, "Automatic modulation recognition using time domain parameters", *Sig. Proc.*, vol. 13, no. 3, pp. 323–328, 1987.

[2] E. E. Azzouz and A. K. Nandi, "Automatic identification of digital modulation types", *Sig. Proc.*, vol. 47, pp. 55–69, 1995.

[3] S. L. Marple, *Digital Spectral Analysis*. Englewood Cliffs, NJ: Prentice Hall, 1987, (Chapter 7).

[4] T. H. Nguyen, M. Sugeno, R. Tong, and R. R. Yager, *Theoretical Aspects of Fuzzy Control*. New York: Wiley, 1995.

[5] A. Papouli, *Probability, Random Variables, and Stochastic Processes*. New York: McGraw Hill, 1991.

[6] S. S. Soliman and S. Z. Hsue, "Signal classification using statistical moments", *Trans. Commun.*, vol. 40, no. 5, pp. 908–916, 1992.

Jerzy Łopatka received the M.Sc. and Ph.D. degrees in communications from the Military University of Technology (MUT), Warsaw. At present he is a Head of Radio and Electronic Warfare Section in the Communication Systems Institute (MUT). His main research interests include digital signal processing in wireless systems.
e-mail: jlopatka@wel.wat.waw.pl
Communication Systems Institute
Military University of Technology
Kaliskiego st 2
00-908 Warsaw, Poland

Maciej Pędzisz received the M.Sc. degree in communications from the Military University of Technology (MUT), Warsaw. At present he is an Assistant in Radio and Electronic Warfare Section in the Communication Systems Institute (MUT). His main research interests include digital signal processing in wireless systems.
e-mail: mpedzisz@wel.wat.waw.pl
Communication Systems Institute
Military University of Technology
Kaliskiego st 2
00-908 Warsaw, Poland

Paper

Modelling of CDMA systems

Piotr Gajewski and Jarosław Krygier

Abstract — This paper describes a model of a system with wideband CDMA that is a proposal of user access for future UMTS. This model has been implemented using OPNET tool. The model enables network architecture, radio interface, mobile station motion, call generation, signalling, etc. The examples of simulation results of call and handover blocking probability are also presented in this paper.

Keywords — OPNET, DS-CDMA, cellular systems, network simulation.

1. Introduction

The commercial proliferation of cellular voice and limited data service has created a great demand for mobile communication and computing. Third generation systems, such as UMTS in Europe, CDMA-2000 in US and TTA I in Korea, are based on combination of integrated fixed and wireless mobile services that form a global personal communication network [11]. Each of them is proposed as a IMT-2000 (international mobile telecommunication) system. Wideband code division multiple-access (WCDMA) has been chosen as the basic radio access technology for such systems. Comparing to the narrow-band CDMA, the WCDMA radio interface offers significant improvement, especially for higher rate and multimedia services.

Recently, different components of CDMA systems as well as theirs behaviour and particular performance have been described in many papers [1, 2, 10, 11]. Most of them have used simulation models as a basic way to resolve many complex problems. There are two approaches to simulate the overall performance of DS-WCDMA system [4]. One is a combined approach where the link level and cellular network level simulations are combined into one package. Another approach is to separate the link and system level simulations to reduce the complexity of the simulators. In both cases specialised simulators are needed.

In this paper, the model of WCDMA system with direct sequence spreading scheme is presented. It was elaborated using MIL-3 OPNET radio-modeller tool. The model consists of network architecture, radio interface, subscribes mobility, traffic generation, channel allocation, handover, signalling, etc. It enables the comprehensive analysis of the influence of many WCDMA system components on the quality of service (QoS). Results of such analysis can be utilised to solve the questions concerning base stations deployment in a given area, optimisation of channel allocation and handover procedures in urban and suburban environments, etc.

2. Simulation tool characteristic

Taking into account the growing interest for WCDMA system modelling, we decided to elaborate our model using MIL-3 OPNET. Optimised network engineering tool (OPNET) is an example of commercial software package that is capable of simulation of large communications networks with detailed protocol modelling and performance analysis. It consists of specialised modules for creation of network and node models, elaboration of models processing, simulation executing, and simulation data analysis as well as output data edition.

OPNET simulation bases on a discrete-event modelling approach, where the progression of the model over simulation time is decomposed into separated time points in which the system state can change.

3. Network architecture

The cellular network is a complex system that includes architecture, procedures and services both on user, network and management levels. The network architecture consists of the mobile stations (MS), base stations (BS), base station controllers (BSC), mobile switching centres (MSC), management centre, and exterior public switching telecommunication network (PSTN) (Fig. 1). These elements are interconnected by transmission links – wired or wireless.

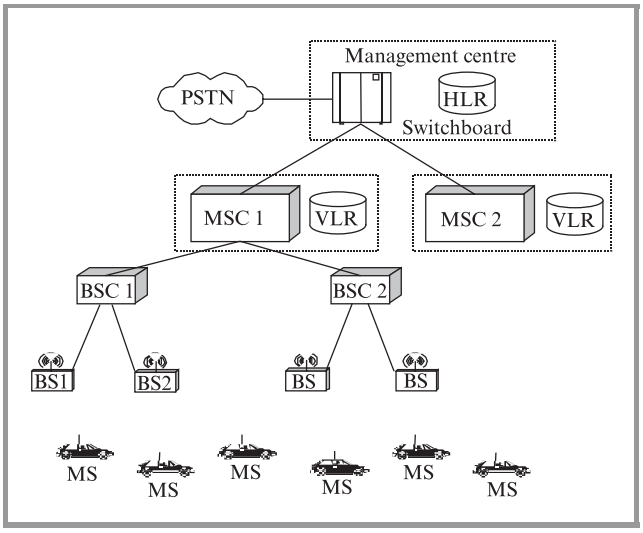


Fig. 1. Cellular network architecture.

The radio link is the duplex communication channel from the base station to each mobile user and from user to base station.

4. Simulation model structure

General simulation model structure is displayed in Fig. 2. The model consists of five basic levels:

- input data setting,
- net (configuration, cell sizes, dimension of switching regions),
- user's behaviour (generation of calls, mobility, soft handover effect),
- call service (allocation and reallocation of channels, transfers, supervising of connections),
- collecting simulation effects.

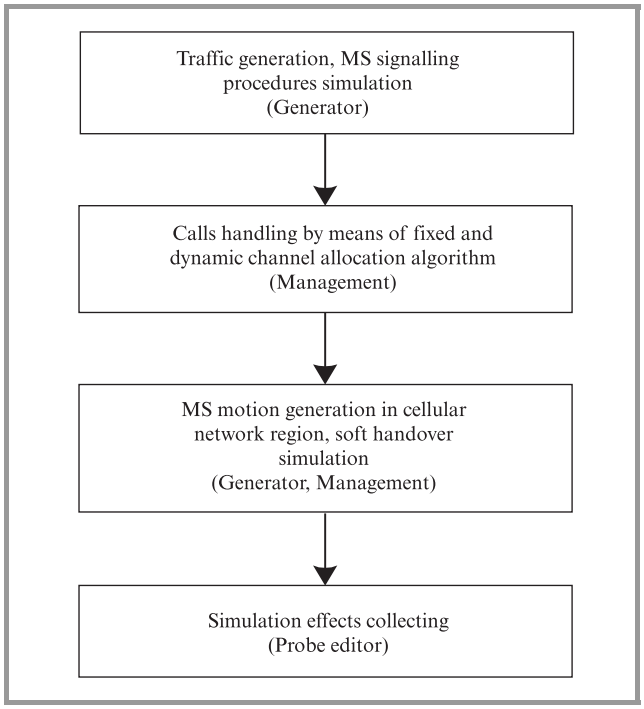


Fig. 2. Simulation program structure.

The following parameters can be determined in the input data set:

- mean number of mobile station users for one cell,
- general amount of duplex channels for speaking,
- amount of fixed and dynamic channels for one cell,
- number of channels reserved for switched calls,
- channel holding mean time,
- mean interval time for successive calls coming from any free subscriber,
- probability of outer subscriber's inaccessibility,
- mean value and standard deviation of mobile velocity,
- cellular network modulus and switching region radius.

5. Real-time services traffic model

In the real-time services case the traffic model is a traditional birth-death process. Speech users arrive to the system according to simple Poisson process with intensity λ:

$$P(k,t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t},$$

where: $P(k,t)$ – probability that there k calls take place during period of time $(0,t)$, t – time, k – number of calls during period of time $(0,t)$.

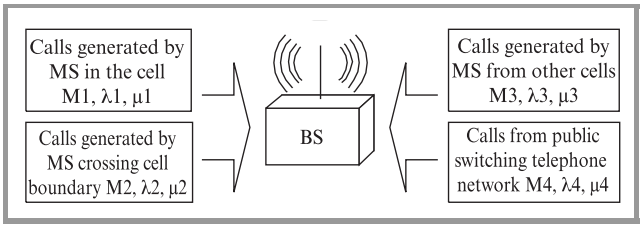


Fig. 3. Types of calls directed to the cell (M1, 2, 3, 4 – number of users, λ1, 2, 3, 4 – calls intensity, μ1, 2, 3, 4 – service intensity).

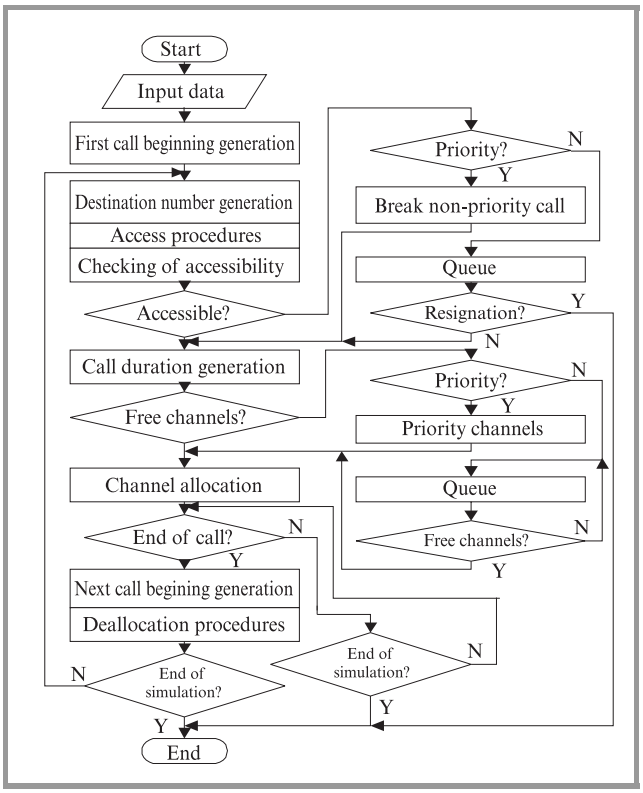


Fig. 4. Traffic generation and calls service algorithm.

The call generators as well as communications traffic procedures were elaborated in presented model. The traffic in each cell consists of four kinds of calls (Fig. 3). They are:

– intra-cell calls generated by MS in the cell that are directed to the other MS in the same cell,

– calls incoming from other cells,

– calls incoming from PSTN,

– calls handovered from neighbouring cells.

Number of users M, calls intensity λ and service intensity μ are the basic parameters of each part of traffic that should be serviced in each cell. Call duration and time between calls are exponential random variables that finally define the traffic intensity generated in cell. Moreover, we assume a number of call priorities in our model.
In case of inter-call, two traffic channels have to be assigned for it. In the other cases, one channel is used for connection. Simplified algorithm of mobile traffic generation is shown in Fig. 4.
It should be noticed, that a call handovered from neighbouring cell has the highest level of priority and that a call generated by PSTN is directed to the cell in which MS is being in a given moment.
The remaining, non real-time services are not taken into consideration in this paper since their models are being elaborated. Channel allocation and deallocation model mention in Fig. 4 is based on modified dynamic channel MDCA allocation policy described in [6].

6. Mobility model

Mobility modelling is involved with the analysis aspects related to location management (location area planning, paging strategies), radio resource management (access technique, channel allocation strategies, handover rates), network signalling loads and propagation (handover decision). Different purposes require different types of mobility models. In vehicular outdoor area, there should be used different mobility model then in indoor environment. Let assume outdoor pedestrian environment, where deployment area is a regular grid of streets and buildings. In this connection the mobile station mobility can be model as fourth-directional motion with the same probability each one. There can be defined motion attributes such as: mean mobile velocity, mean way length, mean time to point of direction change and PDF of applied random variables. The mobile's position is updated every specific time that is dependent on its speed. In such points, remaining call time is checked according to drawn channel holding time.

7. Mobility management model

Mobility management contains two components: handover management and location management [11]. Handover management enables the network to maintain a user's connection as the mobile station to move and change its access point to the network. Handover management includes

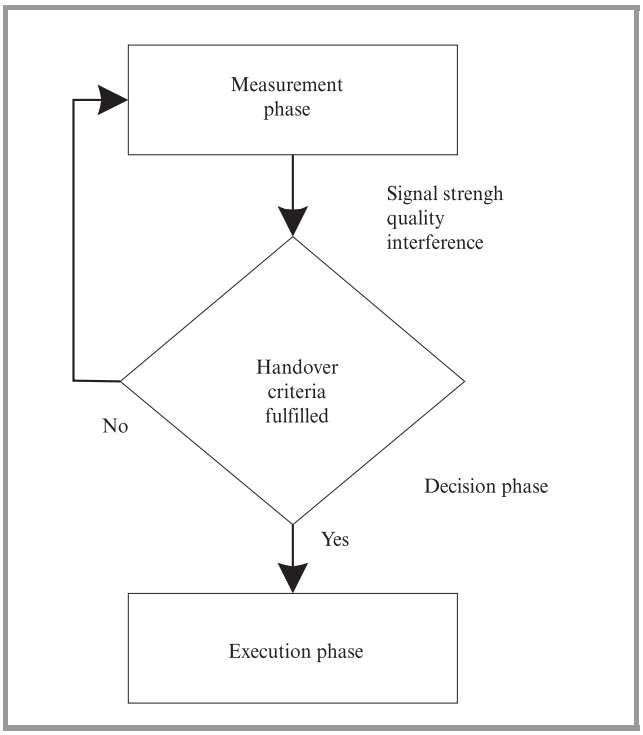


Fig. 5. Handover phases.

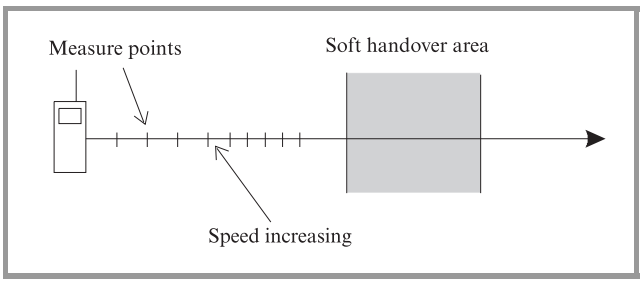


Fig. 6. Measure points and soft handover area.

two conditions intercell and intracell handover. Let assume that we deal with omni directional antennas and only intercell handover will be taken into consideration. A handover procedure in such case can be divided into three phases: measurement, decision and execution phase, as illustrated in Fig. 5 [4].
Due to specific properties of CDMA system connected with the possibility of using the same frequency band in the whole network, the soft handover is often utilised. It means that two or more BS supervise the quality of radio connection in case of the mobile station moving inside the so-called soft handover range and choose the most convenient base station through which useful information exchange is carried on. In wideband CDMA system with asymmetric traffic, at least the following parameters can be identified:

– distance attenuation,

– uplink interference,

– downlink interference.

Built handover model takes into account distance attenuation. The same measures as in mobility model are needed to gather information for a handover decision in such situation. The averaging period for the measurement results depends on the mobile speed and is updated more frequently when the speed of the mobile increases (Fig. 6).

Location management is a two-stage process that enables the network to discover the current attachment point of the mobile user for call delivery. The thirst stage is location registration (location update) and the second one is call delivery. In the location update stage the mobile periodically notifies the network of its place of stay.

The model of location management consists of the way of location area LA as well as base stations distinguishing. All indispensable data are stored in home location register (HLR) and visitor location registers (VLR). BSs, BSCs, MSCs and PLMN are addressed by means of their nametags specified directly in network editor.

They are next converted on ISDN numbers as follows: 10000000*PLMN nametag + 100000*MSC nametag + + 1000*BSC nametag + BS nametag.

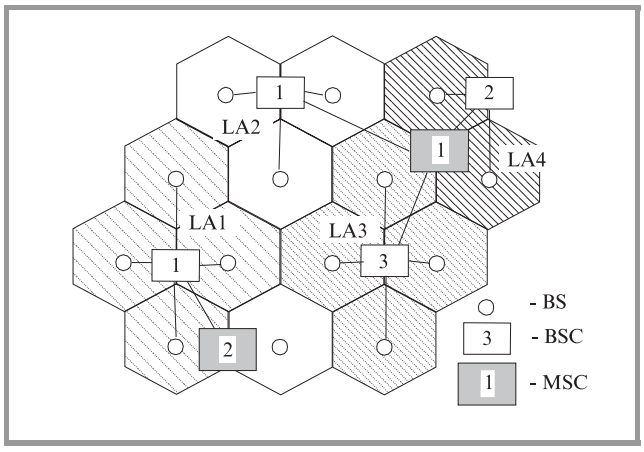


Fig. 7. Modelled location areas.

So the last three digits represent adequate BS, next four digits – BSC and MSC and the last one – PLMN. For example 10302001 is a number of BS with nametag amount to 1, which is connected to BSC with nametag amount to 2, while MSC nametag is amount to 3. VLRs are connected to MSCs, so the number of MSC unambiguously identifies them. Whole network is divided into location areas that are recognised by the BSC number. Such situation is shown in Fig. 7. Therefore, 10201000 is an address of the LA1 location area, 10101000 – LA2, 10103000 – LA3 and 10102000 – LA4.

Home location register is a some kind of data base storing information about registered MSs. On the beginning of the simulation there is created configuration file describing each mobile station. During simulation there are created working files representing VLRs. They are updated according to mobility and handover model.

8. Example of signalling model

When the user of a MS originates a call, he first enters the called number and possibly additional information with the MS keypad and then depresses the send button. Figure 8

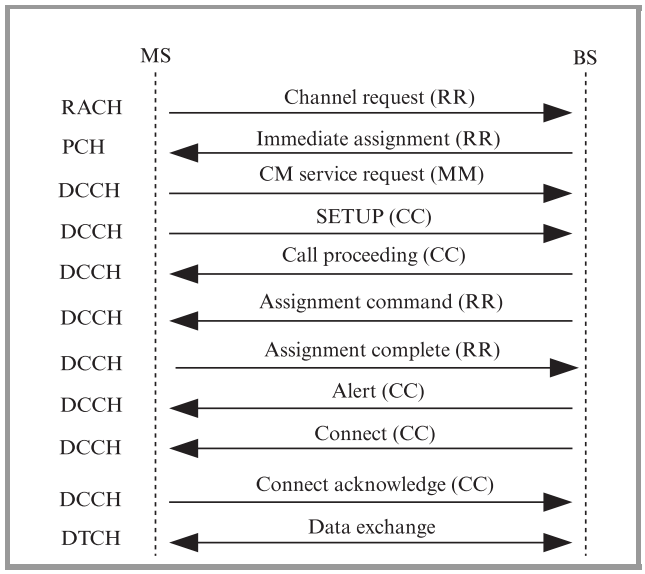


Fig. 8. Signalling for the set-up of the call originated by MS.

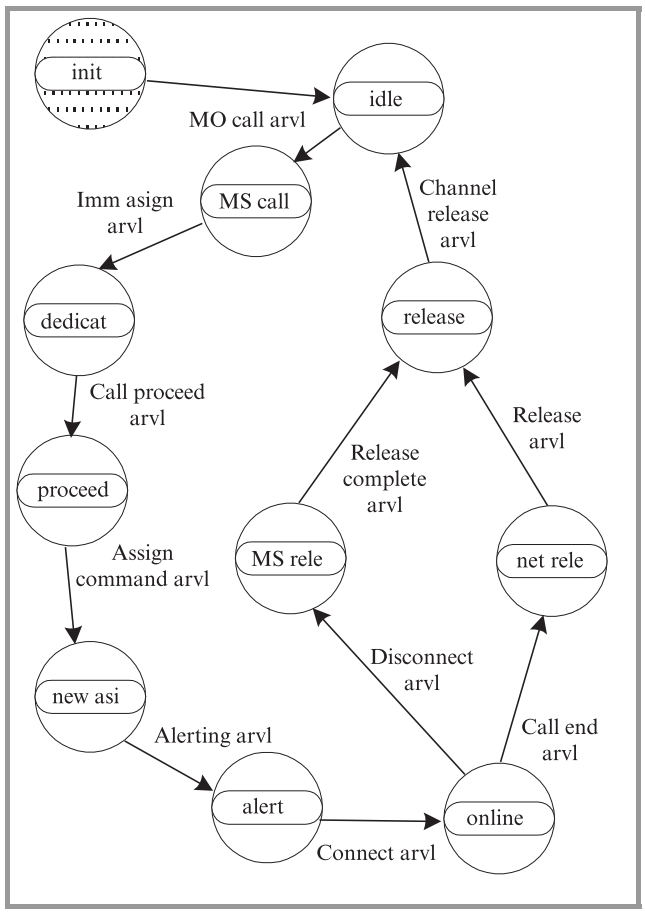


Fig. 9. Example signalling procedure process.

shows the signalling for the set-up of the call which is used in the model.
The abbreviations in Fig. 8 means logical channels: RACH – random access channel, PCH – paging channel, DCCH – dedicated control channel, DTCH – dedicated traffic channel and adequate protocols: RR – radio resource management, MM – mobility management, CC – connection control management.
Such signalling procedures are modelled by their implementation in processes form. The process states and conditional transitions are shown in Fig. 9.
After call beginning, each mobile station model starts its work from "idle" state. Its state is being changed when interrupt named "MO call arv" appears. In the MS call state, mobile is being waited for "Immediate assignment" procedure directed from the network (Fig. 9). After receiving this interrupt it is being moved to dedicated state and send "Request" and "Set-up" procedures to the network. Such situation last as long as the model achieve of "online" state. It means the MS has established the connection.
The MS model state is being came back to the "idle" state after receiving "Call end arvl" or "Disconnect arvl" interrupt. There are two ways of coming back, depending on disconnecting manner (disconnected by MS or by the network). It should be mentioned that this is only a part of whole signalling process model.

9. Example simulation results

The simulation results were presented in [2, 5, 12]. Figures $10 \div 12$ show the example results of investigation of quality of service as a call and handover blocking probability versus mean traffic intensity in the cell for different method of channel allocation, for different values of handover region.
On the basis of presented results of simulation we can notice that:

- usage of the MDCA significantly decreases call blocking probability comparing to fixed channel allocation method,
- handover area decrease causes the call blocking probability increase,
- MS velocity decrease causes the call blocking probability increase.

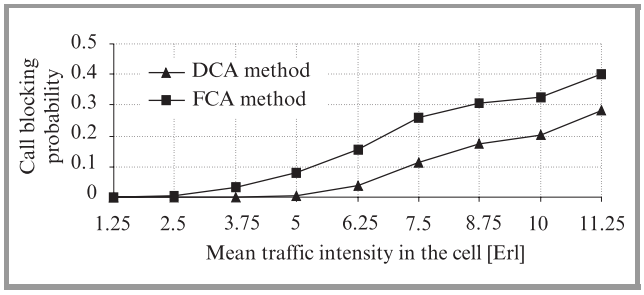


Fig. 10. Call blocking probability versus mean traffic intensity in the cell for two methods of channel allocation.

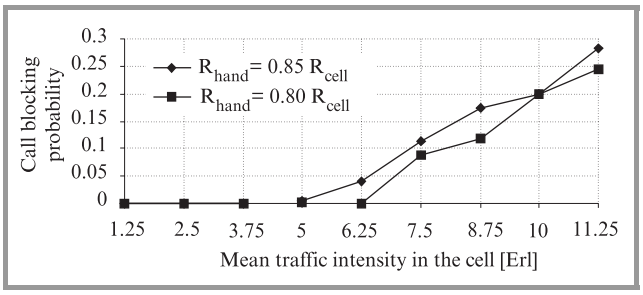


Fig. 11. Call blocking probability versus mean traffic intensity in the cell for two handover radiuses.

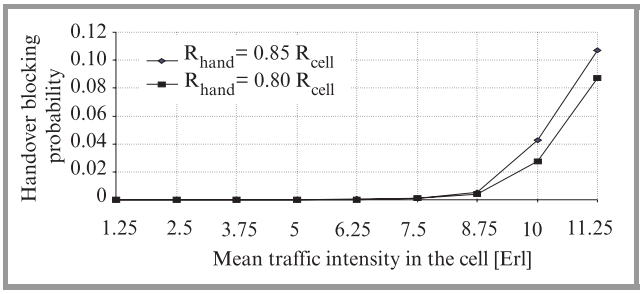


Fig. 12. Handover blocking probability versus mean traffic intensity in the cell for two handover regions.

10. Conclusions

The model of WCDMA system has been elaborated by authors for investigations and planning of universal mobile telecommunications system proposed as IMT-2000 standard. System behaviour was modelled as a set of procedures in OPNET C language. The system properties for modified DCA and soft handover as well as for various methods of channelisation and spreading was investigated. The obtained results confirmed that model is useful and flexible. Model is open and can be extended on the other procedures used in UMTS parts, for example on the information privacy methods.

References

[1] E. Dahlman, P. Beming, J. Knutsson, F. Ovesjo, M. Petrsson, and Ch. Roobol, "WCDMA – the radio interface for future mobile multimedia communications", *IEEE Trans. Veh. Technol.*, vol. VT-47, no. 4, 1995.

[2] P. Gajewski and J. Krygier, "CDMA systems modelling using OPNET software tools", in *Proc. Wirel. Pers. Conf. WPC'98*, Blacksburg, June 1998.

[3] D. E. Everitt, "Traffic engineering of the radio interface for cellular mobile networks", in *Proc. IEEE*, vol. 82, no. 9, 1994.

[4] T. Ojanpera and R. Prasad, *Wideband CDMA for Third Generation Mobile Communications*. London: Artech Hause Publishers, 1998.

[5] P. Gajewski, J. Krygier, J. Łopatka, and J. Buczyński, "Performance of a DS-CDMA system with dynamic channel allocation and soft handover", in *Proc. ISSSTA '98*, Sept. 1998.

[6] E. Del Re, R. Fantacci, and G. Gimbene, "Handover and dynamic channel allocation techniques in mobile cellular networks", *IEEE Trans. Veh. Technol.*, vol. VT-44, no. 2, 1995.

[7] W. C. Lee, "Overview of cellular CDMA", *IEEE Trans. Veh. Technol.*, vol. VT-40, no. 2, 1991.

[8] M. Amanowicz and P. Gajewski, "Compatibility criteria for digital radio systems in nonlinear channels", in *Proc. Africon'96*, Stellenbosch, Sept.-Oct. 1996.

[9] M. B. Pursley and H. F. Roefs, "Numerical evaluation of correlation parameters for optimal binary shift-register sequences", *IEEE Trans. Commun.*, vol. COM-27, no. 10, 1979.

[10] E. K. Hong, K. J. Kim, and K. C. Whang, "Performance evaluation of DS-CDMA system with M-ary orthogonal signaling", *IEEE Trans. Veh. Technol.*, vol. VT-45, no. 1, 1996.

[11] I. F. Akyildiz, J. Mcnair, J. Ho, H. Uzunalioglu, and W. Wang, "Mobility management in next-generation wireless systems", in *Proc. IEEE*, vol. 87, no. 8, 1999.

[12] J. Buczyński, P. Gajewski, and J. Krygier, "Modelling of the third generation mobile system", in *Proc. Africon'99*, Sept. 1999.

Piotr Z. Gajewski received the M.Sc. and Ph.D. degrees from Military University of Technology (MUT), Warsaw, Poland in 1970 and 1979, respectively, both in telecommunication engineering. Since 1970 he works at Electronics Faculty of Military University of Technology (EF MUT) as a scientist and lecturer in communications systems (radios, cellular, microcellular), signal processing, adaptive techniques in communication and communications and information systems interoperability. He was an Associate Professor at Telecommunication System Institute of EF MUT from 1980 to 1990. From 1990 to 1993 he was Deputy Dean of EF MUT. Currently he is the Director of Communications Systems Institute of EF MUT. He is an author (co-author) of over 80 journal publications and conference papers as well as 4 monographs. He is a member of the IEEE Vehicular Technology and Communications Societies. He is also a founder member of the Polish Chapter of Armed Forces Communications and Electronics Association.
e-mail: pgajewski@wel.wat.waw.pl
Military University of Technology
Kaliskiego st 2
01-489 Warsaw, Poland

Jaroslaw Krygier was born in Kolo, Poland in February 15, 1971. In 1991 he joined the Polish Army. He was commissioned as a second lieutenant in 1995. He received the M.Sc. degree from Military University of Technology (MUT), Warsaw, Poland in 1996 in telecommunication engineering. His first posting was to a battalion in MUT. He was a platoon commander. After one-year training period he was posted to Communications Systems Institute in MUT as an engineer. In 1998 he was promoted to lieutenant and since this time he has worked as an Assistant. His main areas of interest are problems of CIS simulation, wideband CDMA technique, CIS interoperability and telecommunication systems engineering.
e-mail: jkrygier@wel.wat.waw.pl
Military University of Technology
Kaliskiego st 2
01-489 Warsaw, Poland

Paper

Performance of CDMA systems with Walsh and PN coding

Piotr Gajewski and Jerzy Dołowski

Abstract — Wideband CDMA method has been recently investigated for its using in the future communication systems. Such systems capacity depends on many factors, the main is multi-access interference (MAI). MAI is an inherent CDMA property connected to correlation characteristics of codes used to access and to spreading. The influence on CDMA performance of these codes families and parameters is presented in this paper. The standard Gaussian approximation has been used for interference analysis. Bit error rate (BER) on detector output is a measure of interference influence on transmission quality. It can be used for SNR_i evaluation on receiver input for the channel assignment and handover procedures. The results of interference analysis for the pseudo-noise sequences (M-sequences, Gold, de'Bruin and Jennings) as well as for the Walsh orthogonal mapping are presented in this paper.

Keywords — PN spreading, CDMA, multi-access interference.

1. Introduction

Recently, wideband code division multiple access (WCDMA) has become the most promising technology for the future personal communication systems (PCS) [1, 2]. This access scheme has a very good behaviour in the urban areas also in case of coexistence with the other cellular systems.

In DS-CDMA systems, channels are created using the specific code sequences. These sequences are used for the spectrum spreading, synchronisation as well as for information transmission. The number of logical channels used is limited by a number of accessible codes in suitable family and by some level of interference. In known standards (IS-95, IS-665), each base station uses the same sets of the code sequences but with different phase shifts [3]. In CDMA standard, the same channel can be used at neighbouring cells; it gives the channel reuse factor approximately equal to one. So, the system capacity is limited by required level of interference from other users from the current and neighbouring cells [4].

Additionally in CDMA, interference results from the non-zero correlation value between two various sequences of used code. The analysis of parameters of codes and their influence on the CDMA performance make possible a choice of codes for the CDMA system. The channel assignment to the particular call could be realised by calculation of the minimal correlation factors with regard to each channel in the interference area.

Thus, the main goal of this paper is to present an evaluation of performance of the DS-CDMA system with different pseudo-noise (PN) as well as orthogonal Walsh code families. Bit error rate BER on detector output is the measure of interference influence on the transmission quality. It can be used for SNR_i evaluation on a receiver input for the channel assignment and handover procedures. BER is defined here as:

$$p_e = \lim_{n\to\infty} \frac{n_i}{n}\,, \tag{1}$$

where: n_i – number of errors, n – total number of transmitted bits.

The results of interference analysis for the pseudo-noise sequences (M-sequences, Gold, de'Bruin and Jennings) as well as for the Walsh orthogonal mapping are presented here. The standard Gaussian approximation was used for BER evaluation. BER was estimated for $n = 10^3 \div 10^5$.

2. Multi-access interference in WCDMA with PN spreading

Let consider the simplified CDMA system with BPSK, with the ideal power control and without the synchronisation errors. We assume a radio channel with $n(t)$ – additive white Gaussian noise (AWGN). Let K is a number of users each generating series $b_k(t)$ of binary data with bit duration T (Fig. 1). Each user's data is multiplied by chip that is a binary PN spreading sequence $a_k(t)$ with duration $T_C \ll T$:

$$a_k(t) = \sum_{j=-\infty}^{\infty} a_j^{(k)}\, p_{T_C}(t - jT_C)\,, \tag{2}$$

where $p_T(t) = 1$ for $0 \le t \le T$ or $p_T(t) = 0$ otherwise.

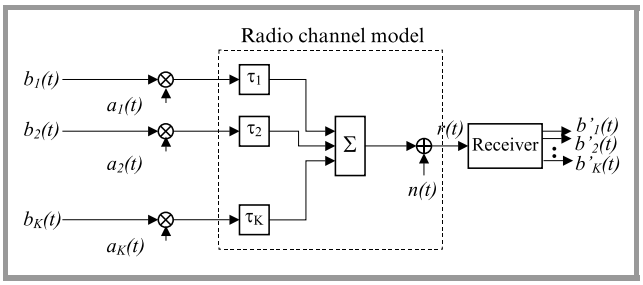


Fig. 1. DS-CDMA model.

We assume that $N = T/T_C$, so we have N chips $a_0^{(k)}, a_1^{(k)}, \ldots, a_{N-1}^{(k)}$ in each data bit.

The signals on the transceiver output and on the receiver input are respectively:

$$s_k(t) = \sqrt{2P}\cdot a_k(t)\, b_k(t)\, \cos(\omega_C t + \theta_k)\,, \tag{3}$$

$$r(t) = n(t) + \sum_{k=1}^{K} \sqrt{2P}\, a_k(t-\tau_k) b_k(t-\tau_k)\cdot\cos(\omega_C t + \phi_k)\,, \tag{4}$$

where: P – power, ω_C – carrier frequency, $\phi_k = \theta_k - \omega_C \tau_k$, θ_k – phase, τ_k – delay of k-th signal.
It can be noticed that $\theta_i = 0$ and $\tau_i = 0$, so $0 < \tau_k < T$ and $0 < \theta_k < 2\pi$ for $k \neq i$.
A signal on the output of the receiver correlator can be expressed as [5]:

$$Z_i = \sqrt{\frac{P}{2}}\left\{ b_{i,0} T + \sum_{\substack{k=1 \\ k\neq i}}^{K} \left[b_{k,-1} R_{k,i}(\tau_k) + \right.\right.$$

$$\left.\left. + b_{k,0} \hat{R}_{k,i}(\tau_k) \right] \cdot \cos\phi_k \right\} + \int_0^T n(t)\cdot a_i(t)\cos\omega_C t dt\,, \tag{5}$$

where $R_{k,i}$, $\hat{R}_{k,i}$ are the correlation functions. For $0 \le lT_C \le \tau \le (l+1)T_C \le T$, they are given by:

$$R_{k,i}(\tau) = C_{k,i}(l-N)T_C + \left[C_{k,i}(l+1-N) + \right.$$

$$\left. - C_{k,i}(l-N) \right] \cdot (\tau - lT_C)\,, \tag{6}$$

$$\hat{R}_{k,i}(\tau) = C_{k,i}(l)T_C + \left[C_{k,i}(l+1) + \right.$$

$$\left. - C_{k,i}(l) \right] \cdot (\tau - lT_C)\,, \tag{7}$$

where $C_{k,i}$ is the aperiodic cross-correlation function for sequences $(a_j^{(k)})$ and $(a_j^{(i)})$ defined as:

$$C_{k,i}(l) = \begin{cases} \sum_{j=0}^{N-1-l} a_j^{(k)} \cdot a_{j+1}^{(i)} & 0 \le l \le N-1 \\ \sum_{j=0}^{N-1+l} a_{j-l}^{(k)} \cdot a_j^{(i)} & 1-N \le l < 0 \\ 0 & |l| \ge N \end{cases}. \tag{8}$$

It is easy to see, the correlation performance of the PN codes should be analysed for MAI evaluation. Signal-to-noise ratio on I-th output of correlation receiver can be estimated using Gaussian approximation taking into account that binary data, time delay and phase shift of each signal are independent random variables. So, Z_i is normally distributed with the mean value $\sqrt{P/2}\cdot T$ and covariance:

$$\sigma^2 = \frac{PT^2}{12N^3} \sum_{\substack{k=1 \\ k\neq i}}^{K} r_{k,i} + \frac{N_0 T}{4}\,, \tag{9}$$

where $r_{k,i}$ can be calculated from [6]:

$$r_{k,i} = \sum_{l=1-N}^{N-1} C_k(l)\left[2C_i(l) + C_i(l+1)\right]. \tag{10}$$

Here, $C_{k,l}$ is the autocorrelation function.

SNR is defined as:

$$SNR_i = \frac{\sqrt{P/2}\cdot T}{\sigma_i}\,. \tag{11}$$

Replacing (9) in (11) the SNR on the output of correlator can be calculated as:

$$SNR_i = \frac{1}{\sqrt{\frac{1}{6N^3}\cdot \sum\limits_{\substack{k=1 \\ k\neq i}}^{K} r_{k,i} + \frac{N_0}{2E}}}\,, \tag{12}$$

where $E = P\cdot T$ is the bit energy.
Thus BER can be estimated from

$$P_e \cong Q(SNR_i)\,, \tag{13}$$

where $Q(\cdot)$ is the error function defined as follows:

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{t^2}{2}}\, dt\,. \tag{14}$$

3. Multi-access interference in WCDMA with Walsh mapping

One of the main methods used for MAI reduction is the orthogonal mapping of data stream generated by the user in the uplink. Each group of data are replaced by the orthogonal sequence in this method. Most often, Walsh sequences are used in practise [1]. The model of quadrature modulator for system with M-ary mapping is presented in Fig. 2.

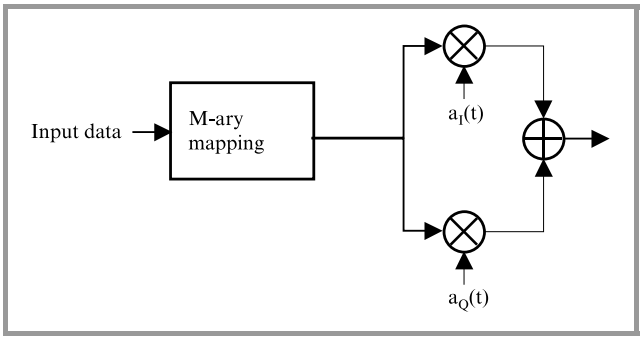


Fig. 2. The quadrature modulator with M-ary mapping.

The orthogonal sequence of k-th user is given by [8]:

$$W^k(t) = \sum_{v=-\infty}^{\infty} W_v^k(t)\cdot P_T(t - vT)\,, \tag{15}$$

where $P_T(t) = 1$ for $0 \le t \le T$ or $P_T(t) = 0$ otherwise.

Here, each of m bits is replaced by one of M Walsh functions of length T (Fig. 3). In (15) W_v^k is the Walsh symbol of k-th user from the set $\{W_0(t), W_1(t), W_2(t), \ldots, W_{M-2}(t), W_{M-1}(t)\}$ corresponding to m bits of input data. Symbol W_v^k consists of M

Walsh "chips" $w_{i,j}\,(j = 0, 1, \ldots, M-1)$ of duration T_W $(T = M \cdot T_W)$.

Therefore

$$W_i(t) = \sum_{j=0}^{M-1} w_{i,j} \cdot P_{T_W}(t - jT_W)\,. \tag{16}$$

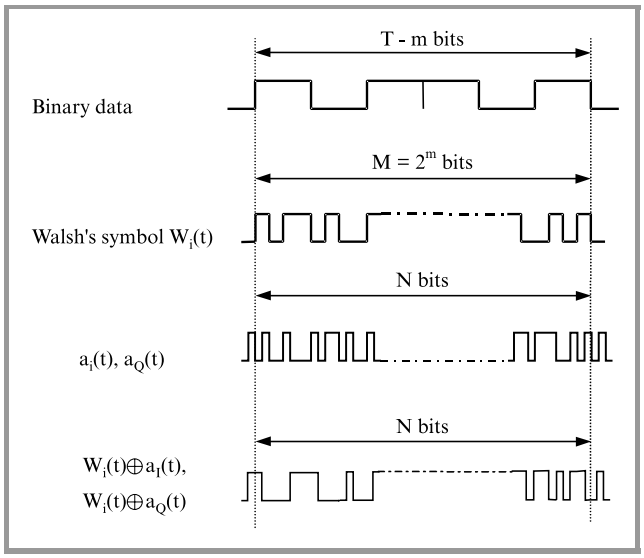


Fig. 3. The sequences in modulator from Fig. 2.

The transmitted signal of k-th user can be expressed as

$$s_k(t) = \sqrt{P} \cdot W^{(k)}(t) \cdot A^{(k)}(t) \cdot \exp(j\theta_k)\,, \tag{17}$$

where $A^{(k)}(t)$ is a complex spreading sequence of k-th user that consists two sequences $a_I^{(k)}$ and $a_Q^{(k)}$ in the in-phase I and quadrature Q branches, respectively.

For a system with K users working asynchronously and simultaneously, the receiver input signal is [8]:

$$r(t) = \sum_{k=1}^{K} \sqrt{P} \cdot W^{(k)}(t-\tau_k) \cdot A^{(k)}(t-\tau_k) \cdot \exp(j\phi_k) + n(t)\,. \tag{18}$$

The detection could be performed using a set of M matched filters [8].

From [8], the variance of signal on the output of k-th filter is

$$\sigma^2 = \frac{1}{2}\mathrm{var}\{Z_k\} = \frac{PT}{24N^3} \sum_{\substack{k=1 \\ k \neq i}}^{K} \gamma_{k,i} + \frac{N_0}{2} \tag{19}$$

and the probability of bit error can be obtained from:

$$p_e = \frac{2^{K-1}}{2^K - 1}\left\{1 - \sum_{k=0}^{M-1} \frac{(-1)^k}{k+1} \cdot \binom{M-1}{k} \cdot e^{-\frac{s^2 k}{2\sigma^2(k+1)}}\right\}, \tag{20}$$

where

$$p_c = \frac{1}{2\sigma^2}\int_0^{\infty}\left(1 - e^{-x/2\sigma^2}\right) \cdot e^{-(x+s^2)/2\sigma^2} \cdot I_0\left(\frac{s\sqrt{x}}{\sigma^2}\right) dx. \tag{21}$$

In (21), $s^2 = P \cdot T$ and $I_0(\cdot)$ is the modified zero-order Bessel function.

4. Numerical results

Basing on above analyses, the simulation program was elaborated in C++. The generators of maximal length sequences (M-sequences), Gold sequences, de'Bruin sequences, Jennings sequences as well as Walsh (Walsh-Hadamard) orthogonal sequences were implemented in this program. It enabled to examine the *BER* as a function of E_b/N_0 for code families mentioned above. Some numerical results are presented below.

Figure 4 shows simulation results for sequences length $N = 63$ (64) and $N = 511$ (512) for $K = 10$ users. The numbers in round brackets concern de'Bruin and Jennings sequences. The length increase causes the *BER* decrease. In practise, the codes of bigger length are used. For example, the codes with $N = 2^{42} - 1$ and $N = 2^{15} - 1$ are used in IS-95 systems.

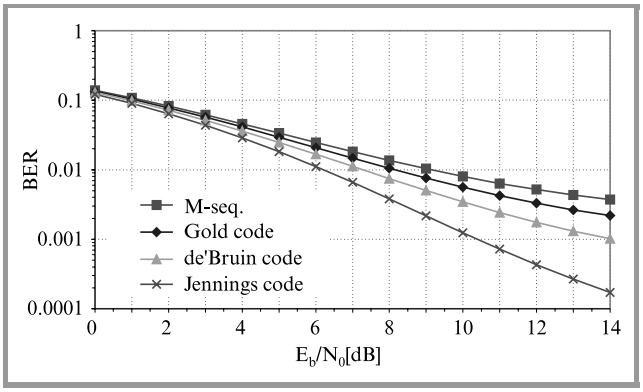


Fig. 4. Simulation results for various codes, $K = 10$, $N = 31$.

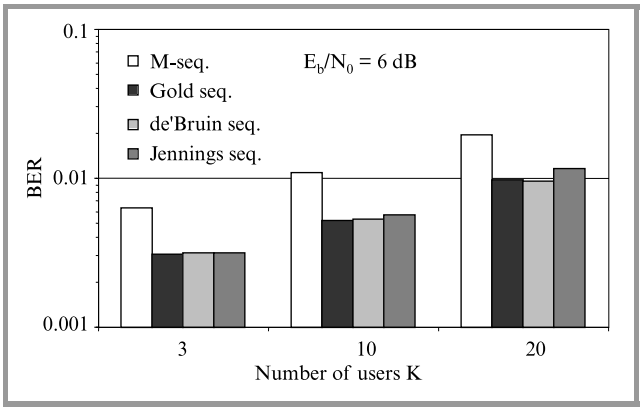


Fig. 5. *BER* versus users number, $N = 127$, $E_b/N_0 = 6$ dB.

Figures 5 and 6 show the *BER* for $K = 3$, 10 and 20 users for code sequences length $N = 127$ (128) and two values of E_b/N_0 equal 6 and 12 dB respectively. It can be observed the differences between results obtained for various codes (except M-sequences) decrease with E_b/N_0 increasing.

Figures 7, 8 and 9 show the results obtained for Walsh mapping with additional PN spreading by Gold and M-sequences. The PN length equals $N = 2^n$, here one 0 bit is added to the sequence for 0 and 1 balancing [8].

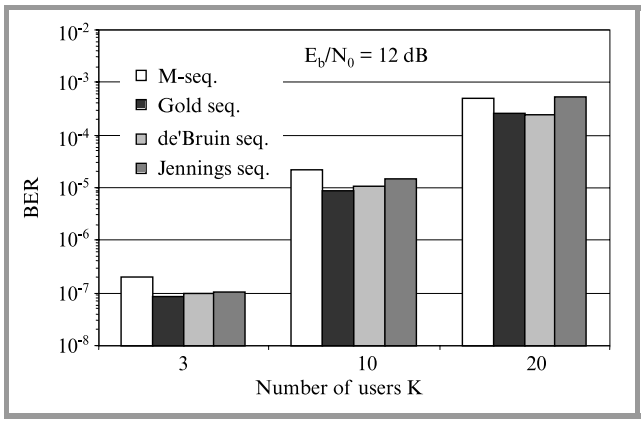


Fig. 6. *BER* versus users number, $N = 127$, $E_b/N_0 = 12$ dB.

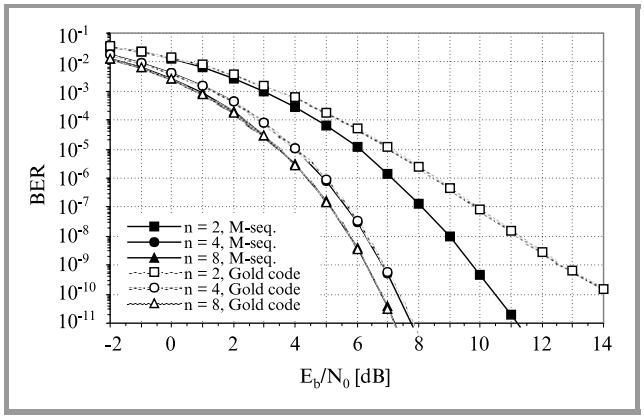


Fig. 7. Simulation results for Walsh mapping, $M = 64$, $K = 10$, added M- and Gold sequences spreading.

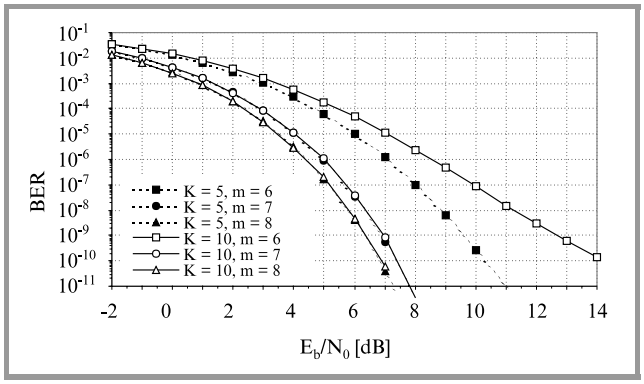


Fig. 8. Simulation results for Walsh sequences, $K = 5, 10$, added Gold sequences spreading.

In Fig. 7, n is the number of PN bits added to one Walsh chip. Here $m = 6$, so one of Walsh 64 bits sequence corresponds to 6 bits of input data. In IS-95, $n = 4$ and $m = 6$. Figure 8 shows the impact of number mapping bits m for two numbers of users $K = 5$ and 10. Here, the $127 + 1$ bits Gold sequence is used for spreading. The m increase decreases the level of MAI.

As it can be seen in Fig. 9, the increase of users number does not significant decrease of the *BER* for longer Walsh sequences. It is caused by *BER* reduction companying the increasing of mapped bits that increases the spreading code length.

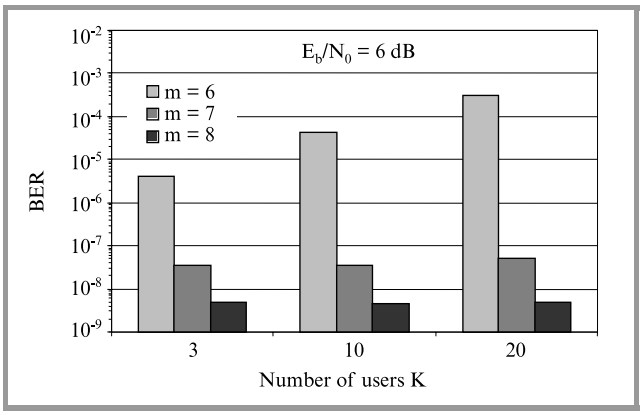


Fig. 9. Simulation results for Walsh sequences, $M = 64$, $K = 3, 10, 20$, $E_b/N_0 = 6$ dB, added Gold sequences $N = 128$.

5. Conclusions

The multi-access is the main source of interference in WCDMA systems.

Taking into account the simulation results it can be concluded that:

- M-sequences have the worst correlation properties comparing to the other considered PN codes; moreover, the number of different sequences in code family is very limited, so the main goal of such codes analysis is they are basis for many other PN codes;
- the very good results was obtained for Gold sequences; it explains these codes usage in real systems;
- the code length has a significant impact on MAI level, so the longer codes are preferred to use in WCDMA systems; that is no problem to generate the long codes, but it can cause an increasing of synchronising time;
- the number of users increasing causes the MAI increasing that limits the system capacity;
- the best results was achieved for Gold and de'Bruin sequences that prefer their usage in WCDMA; but, the performances of Jennings sequences (large number of codes in family, not poor correlation properties, 0–1 symbols symmetry) make these codes very interesting for using in military systems;
- in Walsh mapping, the mapped bits number increase causes the significant reduction of *BER* even a few ranges;
- the very good results in *BER* decreasing are obtained using Walsh mapping companying to PN spreading;

– the PN code type has a less influence on MAI level than a number of mapped bits.

It must be noticed, that presented analysis and results was obtained while some simplified assumptions had been done. It especially concerns the power control in up- and downlinks, users mobility, absence of the noise from narrowband systems, the call traffic performances, etc.
However the results of this analysis can be used in channel access as well as in handover procedures.

References

[1] E. Dahlman, P. Beming, J. Knutsson, F. Ovesjo, M. Petrsson, and Ch. Roobol, "WCDMA – the radio interface for future mobile multimedia communications", *IEEE Trans. Veh. Technol.*, vol. 47, no. 4, 1998.

[2] J. Buczyński, P. Gajewski, and J. Krygier, "Modelling of the third generation mobile system", in *Proc. Africon'99*, Sept. 1999.

[3] P. Gajewski, J. Krygier, J. Łopatka, and J. Buczyński, "Performance of a DS-CDMA system with dynamic channel allocation and soft handover", in *Proc. ISSSTA'98*, Sept. 1998.

[4] M. Amanowicz and P. Gajewski, "Compatibility criteria for digital radio systems in nonlinear channels", in *Proc. Africon'96*, Stellenbosch, Oct. 1996.

[5] M. B. Pursley, "Performance evaluation for phase-coded spread-spectrum multiple access communication". Part I: "Systems analysis", *IEEE Trans. Commun.*, vol. COM-25, no. 8, pp. 795–799, 1977.

[6] M. B. Pursley and D. V. Sarwate, "Performance evaluation for phase-coded spread-spectrum multiple access communication". Part II: "Code sequence analysis" , *IEEE Trans. Commun.*, vol. COM-25, no. 8, pp. 800–803, 1977.

[7] M. B. Pursley and H. F. A. Roefs, "Numerical evaluation of correlation parameters for optional phases of binary shift-register sequences", *IEEE Trans. Commun.*, vol. COM-27, no. 10, 1979.

[8] E. K. Hong, K. J. Kim, and K. C. Whang, "Performance evaluation of DS-CDMA system with M-ary orthogonal signaling", *IEEE Trans. Veh. Technol.*, vol. 45, no. 1, pp. 57–63, 1996.

Jerzy Dołowski received the M.Sc. degree from Military University of Technology (MUT), Warsaw, Poland in 1998 in telecommunication engineering. His first posting was to a radio battalion in communication military unit. He was a platoon commander. After two-year training period he was posted to Communications Systems Institute Electronics Faculty in MUT as an engineer.
e-mail: jdolowski@wel.wat.waw.pl
Military University of Technology
Kaliskiego st 2
01-489 Warsaw, Poland

Piotr Z. Gajewski – for biography, see this issue, p. 40.

Paper

Model of e-book for distance-learning courses

Bogdan A. Galwas, Mariusz Pajer, and Jolanta Chęć

Abstract — Structure of electronic book, prepared as auxiliary materials of courses for specialists in the field of telecommunications and informatics, will be described. Assumptions undertaken and their grounds will be presented. Model of a single lecture, way of a control of student progress and technique of movement through the course content will be described.

Keywords — open and distance learning, continuing education, multimedia, compact disc.

1. Introduction

Continuing education especially professional continuing education have developed for last decade [1, 2], both in respect of the number of courses that are offered and the development of different educational forms. Different tools offered by multimedia techniques and Internet are widely used.

In 1985–95 a synchronous model of distance learning, using satellite television and feed-back for asking the lecturer, was developed. This model is used with pleasure for employee training by corporations, as it gives considerable savings of travel and accommodation (hotels) costs.

Increasing knowledge gained during attending lecturers, their better understanding and getting knowledge how to use the learning are necessity. Thus the possibility of self-education is a very important element in education process.

Traditional tools used for self-education are books, set of lecturers, printing materials. New multimedia technologies and Internet give new possibilities for preparing didactic materials in useful form facilitating material study by learners. Works concerning preparation of a new model of e-book that uses existing possibilities have been launched.

2. Basic assumptions of e-book model

Analysis of needs and possibilities has been made during preparation of the model design of course materials, some assumptions also have been undertaken [3].

First it is assumed that the course material will be provided in electronic form. Because of the big cost of connections and their low quality in Poland it is decided to place didactic material on CD. It will give students the possibility of course content studies without restraint. It is also assumed that material will be published by the use of information service accessible on-line. Location of course content on a single CD limits the amount of material possible for transfer. Though capacity of one disc CD should be enough the possibility of the use of additional discs for content location should be taken into consideration.

Farther it is assumed that the material in electronic form should be transferable between many system platforms. Reproduction (playback) of material should be done by the use of non-commercial software. Tools for review of course content should be provided together with the material.

Next it is assumed that the form of the didactic material should ensure the highest interaction degree for the user. The possibility of most interesting and full presentation of material should be ensured.

Taking into account mentioned above assumptions and experiences of other teams preparing materials for distance learning [3, 4], it is assumed that didactic materials will be developed in the form of HTML documents. Programming language JavaScript will be used for dynamic management of the content of created documents.

It is noticed that authors of didactic materials prefer mostly MS Word editor and use it to prepare documents. It is decided to use MS FrontPage programme to facilitate document conversion from MS Word format to HTML. By the use of MS FrontPage a template of e-book will be developed for material presentation.

Lecture material will be enriched, according to the possibilities, using pictures, animations, sound comments, multimedia presentations, questions and tasks with answers or without answers. Besides basic lecture material, according to the possibilities, auxiliary materials in the PDF documents form (review possible by the use of free Acrobat Reader software) will be provided.

CD-ROM containing lecture and auxiliary materials will be provided with HTML browsers (Netscape Navigator and Internet Explorer), Acrobat Reader and, according to the possibilities and needs, applications for reproduction of multimedia files.

Such material after preparation will be published by the use WWW service accessible through educational portal.

3. E-book structure

The material, with regard to functional aspects, is divided into three main parts: introductory part, lecture's part, auxiliary part.

Introductory part contains large amount of information of different kind but needful and useful. All together is the preparation and introduction for students to appropriate part of material. The content of introductory part may be

different depending on needs. The following elements are assumed to place in this part:

- **Information about authors**. It presents authors of didactic materials, their professional profile, especially taking into consideration professional experience and didactic knowledge.
- **Course purpose**. It is formulated and prepared by authors. It presents level of knowledge and acquirements which students should gain (according to authors intention) after getting knowledge of prepared material. In this part authors can present conditions for understanding didactic material: which minimum level of knowledge and in which areas learners should have to understand the material.
- **Requirements concerning credit for a course**. It is a point specific and characteristic for a given course. They are placed in the case when the course ends with the exam or test.
- **Requirements for equipment** and description concerning the way of CD use. They are a kind of advice and instruction prepared by software developers. Guides placed there may be obvious for many students but for others may be a great assistance.

Lecture's part contains series of didactic units, that can be called lectures. It is assumed that each lecture is a separated unit, basic course quantum, specified entity that student should learn himself. Material of subsequent lectures is obviously connected with each other by the structure and logic of the course and should be studied in the sequence assumed by the author.
Next it was assumed that a lecture is composed according to a given schema and it contains many essential, from student point of view, elements [4]. The elements are as follows:

- **Basic knowledge segments**. They contain didactic material that should be learnt by students. Segments can contain repetitions of material presented during lectures or presented in Internet. Certain parts of material can go beyond the course programme and be an extension recommended to study.
- Questions, problems and tests for **self-evaluation** enable student to get to know the degree of his knowledge. A very important problem is to create the possibility for a student to get to know if his degree of getting knowledge contained in the lecture and understanding of it, meet lecturer expectations and his requirements and in perspective enable to meet examination requirements.
- **Glossary** and subject matter index contain an arrangement of terms and definitions introduced into a given lecture. Glossary enables a student to ensure if all lecture contents are understood and learnt.
- **Bibliography** contains a list of most important items. Getting to know them may be helpful to understand lecture material. They can be as follows book chapters, scientific and technical publications, information contained in Internet, etc.

Understanding of materials and lecturers contents can be facilitated and hastened using auxiliary multimedia tools containing written comments, audio comments, simulations and animations, video insertions, etc. All mentioned, auxiliary multimedia tools should be used, by authors of lectures, purposely and in a well-founded manner. Use of auxiliary multimedia tools differentiates an e-book from a written book. Simple scanning of written-book pages and their presentation on a computer screen does not automatically create the electronic book. Purposeful use of auxiliary multimedia tools creates the new kind of a manual.
Auxiliary part can contain different elements, according to the course subject matter. However, listed below three components should be placed in this part according to the essential role they play in the education process:

- **Index** of all new definitions and terms introduced in subsequent lectures. Each definition and term is provided with a short description and information in which part of material the subject matter was introduced and described. List of that kind facilitates students to remind definitions and terms already known but not yet fix in the memory and eventually their repetition.
- **Library** of materials and publications should not be the usual sum of items mentioned in subsequent lecturers. In this part there is a place to give items, that can be the introduction to the taught material and also items that are the extension of this material. Valuable items can be the texts of current papers from conference proceedings and technical and scientific articles published lately. Lists of Internet addresses to materials placed in Internet by publishers and firms can be also given here.
- **Auxiliary software** attached in this part can facilitate to read certain texts, animations, to do simulation calculations or preparing simple projects.

The most important elements were considered in the assumed model of e-book and their list would probably become longer. However now it can be seen that the number of tools used in such electronic manual exceeds considerably everything that in a paper manual authors can invent and introduce.

4. Pages construction and navigation through the manual

Prepared electronic book was built in the dynamic HTML technology: HTML, Cascading Style Sheets, Java Script.

Thus the whole navigation is based on HTML references and dynamic processing of WWW pages by Java Script. After placing the mouse pointer to the reference the text being a reference is dynamically underlined. Functionally, navigation is a result of a book organisation.

Navigation begins from a front page, it is the **first navigational level**. Information about authors and the first page in a book can be called up from the front page. In each case a new window is opened. The first page of a book is the **second navigational level** and it enables input to single pages, concerning user introduction explaining how to work with a book, or input to the table of contents. Pages containing information about authors and user introduction how to work with a book form the introductory part. The table of contents is the **third navigational level**, it enables to open particular lectures and the index. A lecture is the last and **fourth navigational level**. It is impossible to come from one lecture into another lecture without return to the table of contents. The table of contents and lecturers (containing materials, problems, list of terms and bibliography) form the lecturer's part. Elements of auxiliary part are attainable by dispersed navigation according to its character. Index is accessible from the table of contents and lectures, while library of materials and publications from bibliography.

Construction of a first page of a book (Fig. 1a) is based on three basic elements: an upper information-navigational bar, a left navigational bar and a main window with the essential content of a page. All other pages of a book have the same form. Upper and left bars have azure background and form upside-down letter L. Reference to the previous level of navigation is placed in the left corner of an upper bar. Name of the subject and series of manuals is placed in the middle of an upper bar. A subject title and a page title are placed there (Fig. 1b) in the case of pages containing the user introduction and the table of contents. Left bar on the second navigational level and on pages accessible from it is empty.

Navigation at second and third levels is based on the content of the main window. The first page of a manual contains navigational menu in the form of a list of references to pages with introductory information and the table of contents. The table of contents (Fig. 2a) is a numbered list of lectures with two part description. The number of a lecture is the reference opening a given lecture, name of a lecture is the reference that unroll and roll up list of lecture segments in the table of contents. Number of a segment is the reference that opens a segment in a given lecture. At the bottom of the table of contents reference to the index is placed, that is opened in the new window.

Structure of a lecture (Fig. 2b) is based on the assumption that it is a separated unit, specified entity, that student should get to know. The lecture contains the main contents of a manual. The upper bar informs about the number of a lecture and its title. It contains the reference to the table of contents and **bookmarks** enabling to open each of the lecture functional unit. Bookmarks contain titles of functional parts. The title of a current bookmark is enlarged in comparison with others, colour of background is the same as the colour of background of the essential content (white). Non-active bookmarks have grey background. If pointer is placed under non-active bookmark its background lightens and the title becomes enlarged. In order to come to the part presented by the bookmark, the mouse pointer should be placed under the bookmark and than the left mouse button should be pressed.

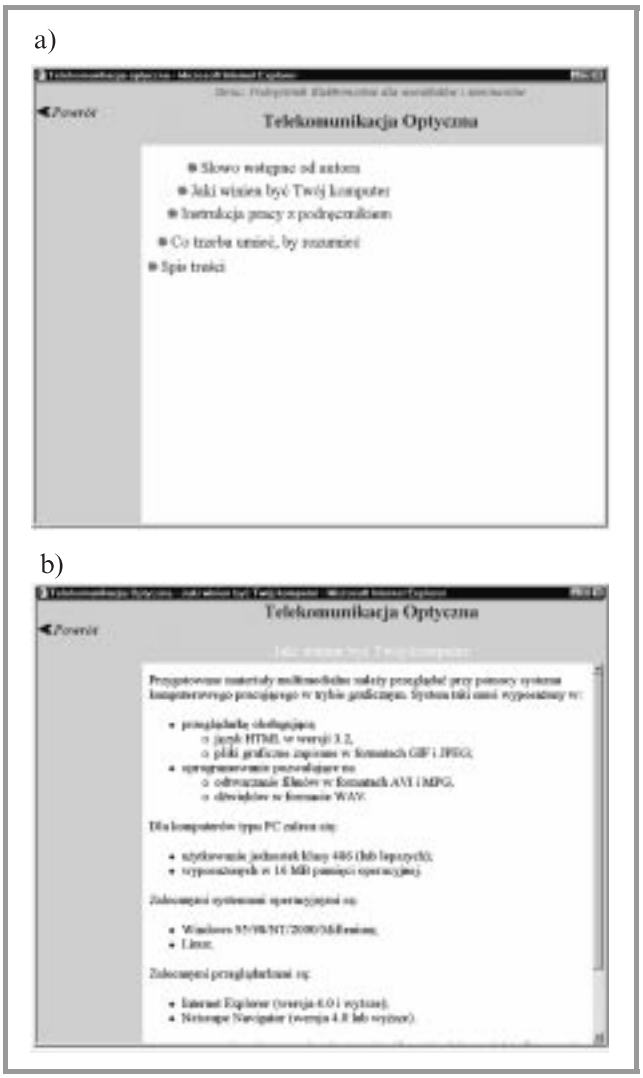


Fig. 1. (a) First page of a book. (b) Description of requirements concerning user computer.

The left bar of WWW page changes according to the actual functional part. Common element for all parts is a **book icon** with a question mark that is the reference to the index. For materials the left bar will contain references to the **segments**. For the current segment references to its **screens** will be specified. In the part "Problems" the left bar can contain the references to particular problems, tasks and tests. For dictionary and bibliography the left bar has not dedicated content. For preparing materials it was assumed that users will have the minimum resolution of a screen

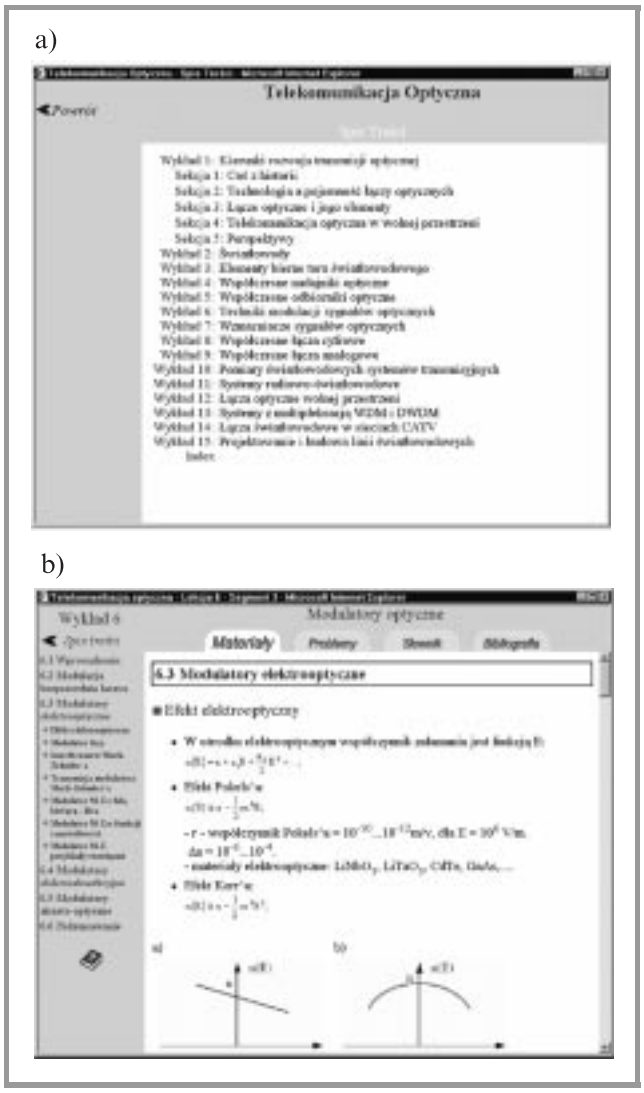


Fig. 2. (a) Table of contents. (b) Lecture material.

800×600 points and they will work in a full-screen mode to acquire materials. The result of the assumption is to assume that the essential content will be presented in a very small degree in a format of 600×500 points. These assumptions are reflected on the organisation of presented material.

The lecture material was divided into segments, for the purpose to acquire a knowledge easier. Each segment is a single HTML file. For the purpose of easier orientation and navigation segments are divided into agreed screens being the smallest basic organisational unit of a lecture material. At the beginning of a page, the title of a segment is placed, farther screens, with screen titles at the beginning, are placed. The segment title and screen titles are HTML references. Important messages are brought into relief by the style or colour of a font. Definitions of laws are denoted by the use of red and blue fonts and the name of a given law is in the bold style. Conclusions are distinguished as a navy blue text on azure background with a key word "Conclusion" in a bold style. Important phrases are brought into relief by the use of a bold style. The signs below figures are written by the use of italics style.

Lecture material is rich in illustrations. Illustrations occur in the different configurations. Single illustrations with a description are the basic form. Pairs of illustrations (arranged horizontally or vertically) with a common sign below also occur. Larger sets of illustrations with a common sign occur seldom. Often the size of readable illustration is too big to be placed on one page. In such case the miniature of illustration is put in the page and it is the reference (distinguished by the use of a colour frame) to the new window filled only with the illustration. Each of illustrations is opened in a separated window. Animations and films are put in four ways: directly on the page (picture with the first frame and control buttons is visible), using reference in the form of a picture with selected frame, using the reference in the form of an icon or using reference in the form of text. Sound elements are accessible by the use of references in the form of an icon with a loudspeaker or reference with text. Items of bibliography can also be references. For the purpose to use the references mouse pointer should be placed on one of them and started by press the left mouse button.

5. Summary

For preparation of a model of an electronic book it was assumed that a student has a contact with the lecturer during the studies. However the contact is not sufficient to get knowledge of the material and a student should additionally by himself work to enlarge his knowledge and its understanding. Electronic book is prepared as a tool for self-education and it should help him in this work.

Periodic contact with a lecturer by the use of Internet and electronic mail can be the assistance in the studies of the material. The contact can be used to send questions and answers, to receive comments concerning tasks and tests being solved.

Compact disc with such prepared material can be used successfully instead of materials printed for courses, seminars and postgraduate studies. It can decrease the number of lecture hours, skipping most of learning process for self-education.

Labour consumption for preparation of an electronic book is much bigger than for a book or sets of lectures written using traditional printing technology. It requires the work of multidisciplinary team and achievement of an appropriate experience. Preparation of single didactic material page in A4 format can take about half or one hour, when there are only simple texts and pictures. It could take three or more hours too, when team have to prepare animations, films, sound commentaries or supplementary applications.

References

[1] B. A. Galwas, "The development of new forms of continuing professional education for telecommunication engineers in Poland", in *7th World Conf. Contin. Eng. Educ.*, Torino, May 10–13, 1998, p. 411.

[2] M. Couzens, "Time-to-knowledge – the new world challenge", in *Proc. 6th Int. Conf. Technol. Supp. Learn. & Train.*, Online Educa, Berlin, 2000, pp. 2–4.

[3] D. Lehman, "Designing hypertext multimedia educational software", *ALN Mag.*, vol. 4, issue 2, Dec. 2000, http//www.aln.org/alnweb/magazine/Vol14-issue2/Lehman.htm

[4] T. Kaskine *et al.*, "The great Paella Cookbook for online learning", Universidad Politecnica de Valencia & Helsinki University of Technology, 1999.

[5] F. Montesinos, J. Lacraz, and C. Montforte, "Case study on CEE change: From face-to-face to open and distance learning CEE", in *Proc. Int. Conf.*, Paris, France, Sept. 2000.

Bogdan A. Galwas was born in Poland, on October 31, 1938. In 1962, he joined the Faculty of Electronics Warsaw University as Lecturer. He received the M.Sc. degree in 1962, the Ph.D. degree in 1969, and the D.Sc. degree in 1976, all in electronic engineering from Warsaw University of Technology, Poland. In 1986 he was promoted to Full Professor. His current research interests are microwave electronics and photonics. He is the author of more than 120 scientific papers and 2 books in these areas. His main field of academic interest is connected with technology of education, continuing engineering education and open distance learning. He is a Chairman of the International Management Committee of the International Travelling Summer Schools '91, Member of IACEE '97 and Member of SEFI '97.
e-mail: galwas@imio.pw.edu.pl
Institute of Microelectronics and Optoelectronics
Warsaw University of Technology
Koszykowa st 75
00-662 Warsaw, Poland

Mariusz Pajer received the M.Sc. Eng. degree in computer engineering from the Warsaw University of Technology (WUT), Warsaw, Poland, in 1999. In 1997 he joined the National Institute of Telecommuniactions (NIT), Warsaw, Poland when he has been working in IT Center in field of data bases, computer networks and WWW. From 1997 he has been working in WUT as network administrator. In 2000 he joined the SPRiNT, Open and Distance Learning (ODL) program, in WUT when he has been working in field of designing and developing course materials and educational portal. From 2000 he has been co-working whith Training Center in NIT in field of ODL. His main interests are information systems, ODL and computer networks.
e-mail: M.Pajer@itl.waw.pl
National Institute of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland

Jolanta Chęć received the M.Sc. degree in electronic engineering (speciality Automatics and Informatics) from the Technical University in Gdańsk. Now she works in the National Institute of Telecommunications (NIT) in Gdańsk (Poland) as a senior specialist. Her research activity is concentrated on information and communication technology (ICT). Since 1994 she has been a member (expert) in the Problem Commission (for Computer Networks and Software) of the Polish Committee for Standardization. In 1995–1996 she took part (research and analytic works) in the international COPERNICUS 1044/STEP Project coordinated by ETSI (European Telecommunications Standards Institute). She is the author of more than twenty published scientific paper (for conferences, symposia, scientific magazines).
e-mail: jolac@il.gda.pl
National Institute of Telecommunications
Jaśkowa Dolina st 15
80-252 Gdańsk, Poland

Paper

IP over optical network: strategy of deployment

Mirosław Klinkowski and Marian Marciniak

Abstract — In this paper, we present main issues of application IP technology directly over optical transport network. The capabilities of WDM transmission systems and techniques which allow the integration of IP with the physical layer are discussed. A detailed description of most solutions is reported. In particular the MPLS/MPλS techniques are discussed in detail. Also problems for further development are outlined.

Keywords — optical networks, WDM, IP, MPLS, MPLambdaS.

1. Introduction

At the rise of the new millennium communication era we witness the convergence of concepts that were quite opposite in the past. Such traditional distinctions as "wired or wireless", "stationary or mobile", "voice or data", or "radio or optical", have lost much of or even all their relevance. This is reflected in novel concepts as "radio over the fibre", "voice over IP, "IP over optical", "free-space optical", "wireless Internet", "convergence of mobile and fixed networks". Further examples are "wireless LANs", "MPLS towards MPλS", or "software defined radio".

There is no doubt that the demand for multi- and hipermedia services foreseen for the near future will require qualitatively new telecommunication technology approaches. Although some of them, e.g. UMTS, 3G, 4G, 5G, or IMT-2000 and beyond, have been already defined, much more should emerge in the near future. Consequently, it is of crucial importance to draw a roadmap for harmonised migration from present communication technologies towards future systems that will be highly efficient at local, regional, and universal scale.

Actually, it is difficult to find a domain of human activity in the world, which would be granted with more global interest and popularity than the Internet and related technologies. Telecommunication technology achievements represented by the exponential growth of the Internet traffic have resulted in the technical availability and the foreseen intense demand for broadband services. Almost all types of traffic are expected to run over Internet. Internet traffic is growing around 10% a month and it is estimated that it would represent over 90 percent of the total public network traffic by the year 2002 [1]. Foreseen immense interest of IP (internet-protocol) technology in the nearest future induces a research for more effective methods of utilising this technology. The most straight and also bringing greatest profit method making possible enlargement of transmission performance seems to be the elimination transmission protocols which lay between physical and IP layers with simultaneous superposition of their functions directly over optical and IP layers.

Development of transmission systems based on dense wavelength division multiplexing (DWDM) technology will make possible building of completely optical networks in the not-distant future where all information will be transmitted and converted as optical signal. Optical crossconects (OXC), optical add-drop multiplexers (OADM) and optical amplifiers will establish optically transparent core of such systems [2]. As it would not be necessary to convert the optical signal to the electrical form for regeneration and switching purposes, the signal degradation resulting from switching procedure as well as the bit-error rate (BER) level will be much lower in those systems [3]. An obvious advantage of DWDM systems is also huge available transmission capacity that significantly exceeds the demand of present applications. At present, transmission systems that can carry up to 40 Gbps on each wavelength are tested commercially [3]. However, the main advantage of DWDM is its compatibility with SONET/SDH, ATM and IP systems.

Two technologies have emerged recently to this purpose. The multi-protocol label switching (MPLS) technology, worked out to assure quality of service (QoS) mechanisms in IP environment, makes possible the integration of IP with other transmission technologies and provides the tools to support different types of services including real-time services [4]. Consequently, the multi-protocol lambda switching (MPλS) technology developed for WDM facilitates wavelength switching mechanism [5].

A huge research is ongoing at global scale to introduce more intelligence in the control plane of the optical transport systems, which should make them more flexible, survivable, controllable and open for traffic engineering. Some of the essential desirable attributes of optical transport networks include real-time provisioning of optical channel trails, providing new capabilities that enhance network survivability, providing functionality and interoperability between vendor-specific optical sub-networks, and enabling protection and restoration capabilities in operational contexts. The research efforts now are focusing on the efficient internetworking of IP with wavelength division multiplexed (WDM) layer.

In the paper the recently proposed technology solutions for IP and optical network integration, grouped according to main issues, are reviewed and discussed.

2. Architecture

One approach for sending IP traffic on WDM networks would use a multi-layered architecture consisting

of IP/MPLS layer over ATM over SONET/SDH over WDM [6]. If an appropriate interface is designed to provide access to the optical network, multiple higher layer protocols can request lightpaths connected across the optical network. An alternative consists of dropping away the ATM layer, and use a packet over SONET/SDH approach, by putting IP/PPP/HDLC into SONET/SDH framing.

The fact that it supports multiple protocols will increase a complexity for IP-WDM integration because of various edge-interworkings required to route, to map, and to protect client signals across WDM subnetworks. The existence of separate optical layer protocols will increase management costs for service providers. Problems could also arise from the high level of multiplexing. The optical fibre transmission systems comprise a large number of higher layer flows such as SONET/SDH, IP flows or ATM VCs. Since these have their own mechanisms, a flooding of alarm messages can take place.

One of the main goals of the integration architecture is to make the optical channel provisioning driven by IP data paths and traffic engineering mechanisms. This will require a tight co-operation of routing and resource management protocols at the two layers. The multi-layered protocol architecture can complicate the timely flow of the possibly large amount of topological and resource information. The two-layer model, which aims at a tighter integration between IP and optical layers, offers a variety of important advantages over the current multi-layer architecture. The benefits include more flexibility in handling higher capacity networks, higher network scalability, more efficient operations and better traffic engineering.

Multi-protocol label switching for IP packets is believed to be the best integrating structure between IP and WDM [7]. MPLS brings two main advantages. First, it can be used as a powerful instrument for traffic engineering. Second, it fits naturally to WDM when wavelengths are used as labels. This extension of the MPLS is called the multi-protocol lambda switching.

Multi-protocol label switching is a switching method in which a label field in the incoming packets is used to determine the next hop. At each hop, the incoming label is replaced by another label that is used at the next hop. It works similarly like switching of the virtual channels (VC) and the virtual paths (VP) in ATM.

The path that is realised this way is called a label switched path (LSP). Each LSP has a set of criteria associated with it, that describe the traffic that the LSP traverses. LSPs can be established using MPLS signalling, e.g., RSVP-TE or CR-LDP [8]. A device that can classify incoming traffic is called a label edge router (LER), while the devices which base their forwarding decision only on the basis of the incoming labels (and ports) are called label switched routers (LSRs).

The model of the network proposed by Internet Engineering Task Force (IETF) consists of both types of IP routers attached to an optical core network (Fig. 1). The optical network consists of multiple optical crossconnects interconnected by optical links. An OXC is a hybrid node consisting of switching element referred to as the optical layer crossconnect (OLXC) and a control plane. Each OXC is capable to switch the data stream from an input port to an output port using a switching function, controlled by appropriately configuring the crossconnect table. OXCs are programmable and should support wavelength conversion and translation.

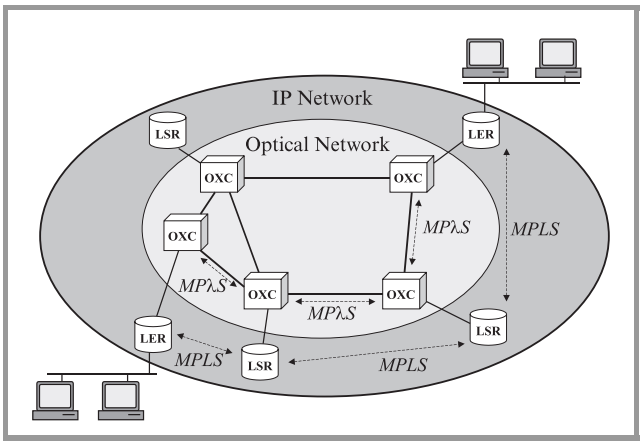


Fig. 1. IP optical network model.

3. Switching in optical network

The exponential growth in Internet traffic leads to a need to scale networks far beyond present speed, capacity and performance. Wavelength division multiplexed transmission and switching are considered the potential solutions to the performance and scaling bottlenecks and offer a limited transparency to packet data-rate and format. However internet-protocol (IP) packet routing and forwarding present a bottleneck in the individual fibre links approach since terabit/s packets must be processed at nearly real-time. Optical WDM switching technology is becoming essential for realisation of the evolving transmission method which is IP over WDM networks.

Evolution of switching in IP/WDM network assumes two phases. First version bases on connection-oriented circuit-switched WDM optical layer, which is supported by MPLS technology. Second phase will make possible fast optical switching in connectionless packet-switched WDM optical layer environment [9]. The two phases are discussed in detail below.

3.1. Phase I

The optical network model consists of multiple optical crossconnects interconnected by optical links in a general topology (referred to as an "optical mesh network"). Each OXC is assumed to be capable of switching a data stream from a given input port to a given output port. This switching function is controlled by appropriately configuring a crossconnect table. Conceptually, the crossconnect

table consists of entries of the form <input port i, output port j>, indicating that data stream entering input port i will be switched to output port j. An "lightpath" from an ingress port in an OXC to an egress port in a remote OXC is established by setting up suitable crossconnects in the ingress, the egress and a set of intermediate OXCs such that a continuous physical path exists from the ingress to the egress port. Lightpaths are assumed to be bi-directional, i.e., the return path from the egress port to the ingress port follows the same path as the forward path.

Multi-protocol label switching has been proposed as the integrating structure between IP and optical layers in the Phase I.

3.2. Phase II

The routing method supports autonomous transmission method: once the packet enters the network, it propagates uninterruptedly to the final destination, excluding the small delay at each routing junction (order of several tens of nanoseconds). In high-performance optical-networks, optical data packets require rapid routing at all-optical switches. In order to achieve an efficient and high-throughput switching, the individual routing bits in the address header must be recovered sufficiently fast.

One of the methods which facilitates a fast conversion of the address header leading the data packet from source to destination was developed by BATM Advanced Communications and Lynx-Photonic-Networks [10]. The method is based on an all-optical-header recognition, whereby the address header, leading the optical data packet, is stripped off the packet, and side tracked to a recognition module (Fig. 2). The method is suitable for address bit rates of 2.5 GHz, 10 GHz and beyond. Since the data block remains in the optical domain, it can be encoded in a possibly high (and practical) bit rate.

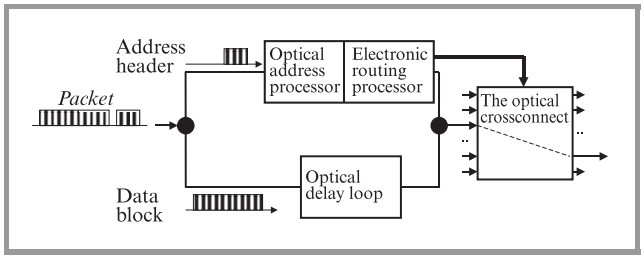


Fig. 2. The routing system based on all-optical address recognition method.

The generated instructions control an OXC, control the entire routing procedure including the regeneration of the new optical address header, restoring it in the proper position. Once the regenerated address has been transmitted into the configured OXC, the data block emerges from the delay loop, following the same path. Each of the OXC's inputs has a dedicated routing module. If collision events (different inputs requiring same OXC output) should become an issue, an optical buffer could be added at the OXC input side, controlled by the electronic module.

4. Signalling & control

Internet Engineering Task Force proposes that control in the IP over optical networks is supported by MPLS control plane. In this solution each node consists of an integrated IP router and reconfigurable optical layer crossconnect equipped with an MPLS control plane.

The IP router is responsible for all non-local management functions, including the management of optical resources, configuration and capacity management, addressing, routing, traffic engineering, topology discovery, exception handling and restoration.

The IP router implements the necessary IP protocols and uses IP for signalling to establish lightpaths. Specifically, optical resource management requires resource availability per link to be propagated, implying link state protocols such as OSPF [7]. On each link within the network, one channel is assigned as the default routed (one hop) lightpath. The routed lightpath provides router to router connectivity over this link. These routed lightpaths reflect (and are thus identical to) the physical topology. The assignment of this default lightpath is by convention, e.g. the "first" channel. All traffic using this lightpath is IP traffic and is forwarded by the router. All control messages are sent in-band on a routed lightpath as regular IP datagrams, potentially mixed with other data but with the highest forwarding priority. Multiple channels on each link are assumed, a fraction of which is reserved at any given time for restoration. The default-routed lightpath is restored on one of these channels. Therefore, we can assume that as long as the link is functional, there is default routed lightpath on that link.

The IP router communicates with the OLXC device through a logical interface. The interface defines a set of basic primitives to configure the OLXC, and to enable the OLXC to convey information to the router.

The mediation device translates the logical primitives to and from the proprietary controls of the OLXC. Ideally, this interface is both explicit and open. One recognises that a particular realisation may integrate the router and the OLXC into a single box and use a proprietary interface implementation. Signalling among various nodes is achieved using CR-LDP and RSVP protocols [7].

Any network addressable element must have an IP address. Typically these elements include each node and every optical link and IP router port. When it is desirable to have the ability to address the individual optical channels those are assigned to IP addresses as well. The IP addresses must be globally unique if the element is globally addressable. Otherwise domain unique addresses suffice. A client must also have an IP address by which he is identified. However, optical lightpaths could potentially be established between devices that do not support IP (i.e., are not IP aware), and consequently do not have IP addresses. This could be handled either by assigning an IP address to the device, or by

assigning an address to the OLXC port to which the device is attached. The network using traditional IP mechanisms can discover whether or not a client is IP aware.

5. Optical network management

Management in a system refers to a set of functions like performance monitoring, link initialisation, and other network diagnostics that allow verifying safe and continued operation of the network. The management functionality in optical networks is still far from being fully developed. The wavelengths in the optical domain will require routing, add/drop, and protection functions that can only be achieved through the implementation of network-wide management and monitoring capabilities.

The links between OXCs will carry a number of user bearer channels and possibly one or more associated control channels. Link management protocol (LMP) that can be run between neighbouring OXCs can be used for both link provisioning and fault isolation [7]. A unique feature of LMP is that it is able to isolate faults independent of the encoding scheme used for the bearer channels. LMP will be used to maintain control channel connectivity, verify bearer channel connectivity, and isolate link, fibre, or channel failures within the optical network.

Current-generation DWDM networks are monitored, managed and protected within the digital domain, using SONET/SDH and its associated support systems. However, to leverage the full potential of wavelength-based networking, the provisioning, switching, management and monitoring functions have to move from the digital to the optical domain. The information generated by the performance monitoring operation can be used to ensure safe operation of the optical network. In addition to verify the service level provided by the network to the user, performance monitoring is also necessary to ensure that the users of the network comply with the requirements that were negotiated between them and the network operator. For example, one function may be to monitor the wavelength and power levels of signals being input to the network to ensure that they meet the requirements imposed by the network. Current performance monitoring in optical networks requires termination of a channel at an optical-electrical-optical conversion point to detect bits related to BER of the payload or frame. However, while those bits indicate if errors have occurred, they do not supply channel-performance data. This makes it very difficult to assess the actual cause of the degraded performance.

Fast and accurate determination of the various performance measures of a wavelength channel implies that measurements have to be done while leaving it in optical format. One possible way of achieving this is by tapping a portion of the optical power from the main channel using a low loss tap of about 1%. In this scenario, the most basic form of monitoring will utilise a power-averaging receiver to detect loss of signal at the optical power tap point. The existing DWDM systems use optical time-domain reflectometres to measure the parameters of the optical links.

The procedure of determining the threshold values for the various parameters at which alarms must be declared causes another problem. Very often those values depend on the bit rate on the channel and should ideally be set up depending on the bit rate. In addition, since a signal is not terminated at an intermediate node, if a particular wavelength fails, all nodes along the downstream path of the failed wavelength could trigger an alarm. This can lead to a large number of alarms for a single failure, and provokes the determination of the cause of the alarm and alarm correlation excessively complicated.

Care should be taken in exchanging those performance parameters. The vast majority of commercial telecommunication networks use framing and data formatting overhead as the means to communicate between network elements and management system. It is worth to mention that while the signalling is used to communicate all monitoring results, the monitoring itself is done on the actual data channel, or some range of bandwidth around the channel. Therefore, all network elements must be able to transmit this bandwidth in order to monitor the events that happen at any point in the network.

6. Fault restoration in optical networks

With the introduction of IP in telecommunications networks there is tremendous focus on reliability and availability of the new IP-optical hybrid infrastructures. Automated establishment and restoration of end to end paths in such networks require standardised signalling and routing mechanisms. Layering models that facilitate fault restoration are discussed. A better integration between IP and optical will provide opportunities to implement a better fault restoration.

Telecom networks have traditionally been designed with rapid fault detection, rapid fault isolation and recovery. With the introduction of IP and WDM in these networks, these features need to be provided in the IP and WDM layers also. Automated establishment and restoration of end-to-end paths in such networks requires standardised signalling, routing, and restoration mechanisms.

The concept of SONET/SDH ring architectures can be extended to WDM self-healing optical rings (SHRs). As in SONET/SDH, WDM SHRs can be either path switched or line switched. Multiwavelength systems add extra complexity to the restoration problem. Under these circumstances, simple ring architecture may be not sufficient.

Although the current routing algorithms are very robust and survivable, the amount of time they take to recover from a failure can be significant, on the order of several seconds or minutes, causing serious disruption of service in the interim. This is unacceptable to many operators that aim to provide a highly reliable service, and thus require recovery times on the order of tens of milliseconds. This problem can be solved by MPLS standards enabled for IP

layer restoration. The MPLS label-switched paths disrupted by a link failure are restored in the IP layer itself, and no restoration occurs in the OLXC layer. MPLS enables a hierarchy of LSPs to be defined by pre-pending a stack of labels or tags to packet headers. When a failure occurs, packets along a given LSP can be routed to a predefined restoration LSP by modifying the labels maps of the routers at the ends of the original LSP [11]. In fact, a protection priority could be used as a differentiating mechanism for premium services that require high reliability.

There are same particular goals for MPLS based protection. First, MPLS-based recovery mechanisms should facilitate fast (within 10 ms) recovery times. MPLS-based recovery techniques should be applicable for protection of traffic at various granularities and for an entire end-to-end path or for segments of an end-to-end path. As an example, it should be possible to specify MPLS-based recovery for a portion of the traffic on an individual path, for all traffic on an individual path, or for all traffic on a group of paths. MPLS-based recovery mechanisms should be able to take into consideration the recovery actions of other layers. MPLS-based recovery mechanisms should also minimise the loss of data and packet reordering during recovery operations.

7. Conclusion

Development of IP, MPLS and WDM technologies makes possible close integration of the Internet applications based on IP technology and optical network infrastructure. Customers will have access to new high-bandwidth services made possible by the increased capacity affordable by the optical layer. Services that today are considered prohibitively expensive, such videoconferencing to the desktop, electronic commerce, and high-speed video imaging, will become commonplace because they will be technologically and economically feasible. It is a giant step forward towards fulfilment idea of information society, which means wide utilisation of information techniques and universal access to communication applications and resources of information.

Acknowledgement

The work has been performed within the framework of European Project COST 266 *Advanced Infrastructure for Photonic Networks*.

References

[1] S. Masud, "Transforming the Network Core", *Telecommun. Online*, Feb. 2000.

[2] M. Marciniak, "Transparency of optical networks: How to manage it?", in *Int. Conf. Transpar. Opt. Netw. ICTON'99*, Th.B.2, Kielce, Poland, 9–11 June 1999, pp. 85–88.

[3] A. Jajszczyk, "Fixed networks evolution", *Telecommun. Rev.*, no. 1, pp. 16–20, 2001.

[4] Ch. Semeria, *Multiprotocol Label Switching (MPLS)*. Juniper Networks, Inc., 2000.

[5] D. O. Awduche, Y. Rekhter, J. Drake, and R. Coltun , "Multi-Protocol Lambda Switching", *IETF Internet draft*, Jan. 2001.

[6] A. Jajszczyk, "Long distance telecommunication networks without ATM and SDH?", *Telecommun. Rev.*, no. 6, 1999.

[7] N. Chandhok, A. Durresi, R. Jagannathan, R. Jain, S. Seetharaman, and K. Vinodkrishnan, "IP over Optical Networks: A Summary of Issues", *IETF Internet draft*, July 2000.

[8] B. Rajagopalan, J. Luciani, D. Awduche, B. Cain, B. Jamoussi, and D. Saha, "IP over Optical Networks: A Framework", *IETF Internet draft*, Nov. 2000.

[9] M. Martin-Villalba, "Research Activities in WDM Routing Network for IP over WDM Networks", July 1999, http://raven.etri.re.kr/semothers/ipwdm/index.htm

[10] R. Krause, "All-optical method for routing of data packets in optical communications networks", 2001, http://www.batm.at/en/press/titan8.htm

[11] R. D. Doverspike, S. J. Philips, and J. R. Westbrook, "Transport network architectures in an IP world", in *Proc. INFOCOM'2000*, Tel-Aviv, March 2000.

Mirosław Klinkowski received his M.Sc. degree from Warsaw University of Technology, Department of Electronics and Information Technology, Poland, in 1999. Actually he is with the National Institute of Telecommunications, Department of Transmission and Fibre Technology. His research interests include telecommunications systems, ATM, SDH, IP over Optical Networks. He is working towards Ph.D. thesis on QoS in IP Optical Networks. He is an author or co-author of 3 papers in international scientific journals and conferences.
e-mail: mklinkow@itl.waw.pl
National Institute of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland

Marian Marciniak, Associate Professor has been graduated in physics from Marie-Curie Sklodovska University in Lublin, Poland, in 1977. From 1985 to 1989 he performed Ph.D. studies in electromagnetic wave theory at the Institute of Fundamental Technological Research, Polish Academy of Sciences, followed by Ph.D. degree (with distinction) in optoelectronics received from Military University of Technology in Warsaw. In 1997 he received his Doctor of Sciences (Habilitation) degree in optics from Warsaw University of Technology. From 1978 to 1997 he held an academic position in the Military Academy of

Telecommunications in Zegrze, Poland. In 1996 he joined the National Institute of Telecommunications in Warsaw where he actually leads the Department of Transmission and Fibre Technology. His research interests include photonics, terabit networks, IP over WDM networks, optical waveguide theory and numerical modelling, beam-propagation methods, and nonlinear optical phenomena. He is an author or co-author of over 160 scientific publications in those fields, including three books. He is an active member of the IEEE – Lasers & Electro-Optic Society, IEEE – Communications Society, New York Academy of Sciences, Optical Society of America, SPIE – The International Society for Optical Engineering and its Technical Group on Optical Networks, and American Association for the Advancement of Science. He was the originator of accession of Poland to European Research Programs in the optical telecommunications domain: COST 240 *Modelling and Measuring of Advanced Photonic Telecommunication Components*, COST 266 *Advanced Infrastructure for Photonic Networks*, COST 268 *Wavelength-Scale Photonic Components for Telecommunications*, and COST P2 *Applications of Nonlinear Optical Phenomena*. He has been appointed to Management Committees of all those Projects as the Delegate of Poland. He has been appointed as the Evaluator of the European Union's 5th Framework Program proposals in the Action Line *All-Optical and Terabit Networks*. He is a Delegate to the International Telecommunication Union, Study Group 15: *Optical and Other Transport Networks*, and to the International Electrotechnical Commission, Technical Committee 86 *Fibre Optics* and its sub-Committees. He served as a Delegate to the *World Telecommunication Standards Assembly* WTSA 2000. He is the originator and the Chairman of the Topical Commission on *Fibre Technology* of the National Committee for Standardisation. In early 2001 he originated the IEEE/LEOS Poland Chapter and actually he serves as the Interim Chairman of that Chapter. He participates in Program Committees of several international conferences, and he is a reviewer for several international scientific journals. In addition to that, he serves as a Member of the Editorial Board of *Microwave & Optoelectronics Technology Letters* journal, Wiley, USA, and the *Journal of Telecommunications and Information Technology*, National Institute of Telecommunications, Poland. He was the originator and the organiser of the 1st, 2nd and 3rd *International Conferences on Transparent Optical Networks ICTON'99*, 2000, and 2001. Recently he has been appointed by the *World Scientific and Engineering Society* to act as the originator and Chairman of the *WSES International Conference on Wireless and Optical Communications* 2002. His biography has been cited in Marquis *Who's Who in the World*, *Who's Who in Science and Engineering*, and in the *International Directory of Distinguished Leadership* of the *American Biographical Institute*.

e-mail: M.Marciniak@itl.waw.pl
National Institute of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland

Paper

Analysis of optical-microwave mixing process in electro-optical modulators

Bogdan A. Galwas and Zenon R. Szczepaniak

Abstract — The principle of operation and the general parameters of the electro-optical Mach-Zehnder modulators are reminded in the paper. With the use of the mathematical relationships describing the transmission of the optical signal through the modulator, the theoretical model of the optical signal transmission with two modulating signals at two frequencies will be presented. The theoretical relationships describing the efficiency of the optical-microwave mixing and frequency multiplying in the two-tone operation will be derived. The analysis and the simulations will be performed for different operating points, where the nonlinearity of the transmission characteristic is specially strong.

Keywords — electro-optical modulator, nonlinear effects, microwave-optical mixing.

1. Introduction

During the last years the nonlinear effects in the processes of optical signals modulation and detection are investigated very intensively. It is related to the development in a new type of optical fibre links with capability of transmitting lots of electrical signals with the use of many carrier frequencies (subcarrier multiplexing) and achieving an operation without the intermodulation distortions.

The analysis of the nonlinear effects allows explaining the electrical-optical mixing processes and further it allows using them for different purposes.

In the photodetection processes the PIN photodiodes and the phototransistors are commonly used. In both cases the nonlinear effects are observed, who could be used in the frequency conversion process. The modulation of the optical signals can be obtained in different ways. When the laser generation conditions are changed in this process, thereby causing the changes in the power or the frequency of the generated signal – it is so-called internal modulation. Whereas when there is no disturbance of the generation process and the optical signal parameters are being changed during the transmission through the separate device placed at the laser output – it is so-called external modulation. There are two types of modulators, which are now commonly used: electro-absorption modulators and electro-optical modulators called Mach-Zehnder modulators. The characteristics of the second ones are well known and described.

The electro-optical modulators are readily used in the analogue optical links, for the sake of exceptional good transmission conditions while operating with the high levels of the optical power. The important problem that appears is the linearization of the transmission characteristics and the minimization of the nonlinear effects. However, the nonlinear effects could be used in the frequency conversion processes for obtaining required conversion products, for example, such as the harmonics of the modulation signal. The work presented here treats of analysis of the conditions to rise up such products.

2. Mach-Zehnder electro-optical modulator

The electro-optical effect appears when the refraction coefficient *n(E)* of an optical material is a function of electric field E

$$n(E) = n + a_1 E + \frac{a_2}{2} E^2 + \ldots \cong n - \frac{1}{2} n^3 r E \,. \qquad (1)$$

When the two of the first parts of the expression (1) are dominant in the electro-optical material, this phenomenon is called Pokels effect. In the expression (1) the Pockels coefficient is denoted as r and it is equal to $10^{-10} \ldots 10^{-12}$ m/V at the electric field $E = 10^6$ V/m, then $\Delta n = 10^{-6} \ldots 10^{-4}$. The materials having this effect are called the electro-optical ones and they are, for example, $LiNbO_3$, $LiTaO_3$, CdTe, GaAs [1].

With the use of the Pockels effect it is possible to make an optical signal phase modulator. In the material demonstrating the electro-optical effect, for example, $LiNbO_3$, a planar optical waveguide is made. This optical waveguide is placed in an electric field created by the electrodes. Then the electric field as a result of a signal connected to the electrodes causes a change of the phase of an optical signal transmitted through the optical waveguide.

The use of an interferometer allows constructing an optical signal amplitude modulator. A planar version of such modulator is shown in Fig. 1. An optical beam with the power P_0 is divided in two equal parts and directed into two branches. Within theses branches two electro-optical phase modulators are placed and they work in counter-phase. The next coupler sums the two beams. When the effects of a high frequency and the attenuation of the planar optical waveguides are neglected, the power transmission of the modulator can be described by a simple formula. The operation of both phase shifters in the interferometer branches is shown in Fig. 2a [2].

The output signal and further the transmission of the modulator can be obtained by summing the optical signals from two branches. It is shown in Fig. 2a.

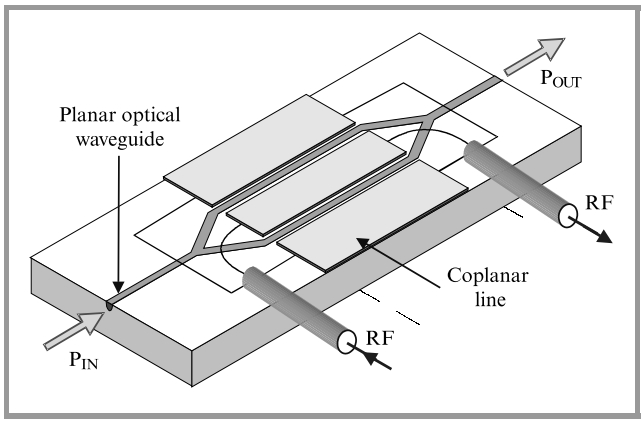


Fig. 1. Planar structure of electro-optical Mach-Zehnder modulator.

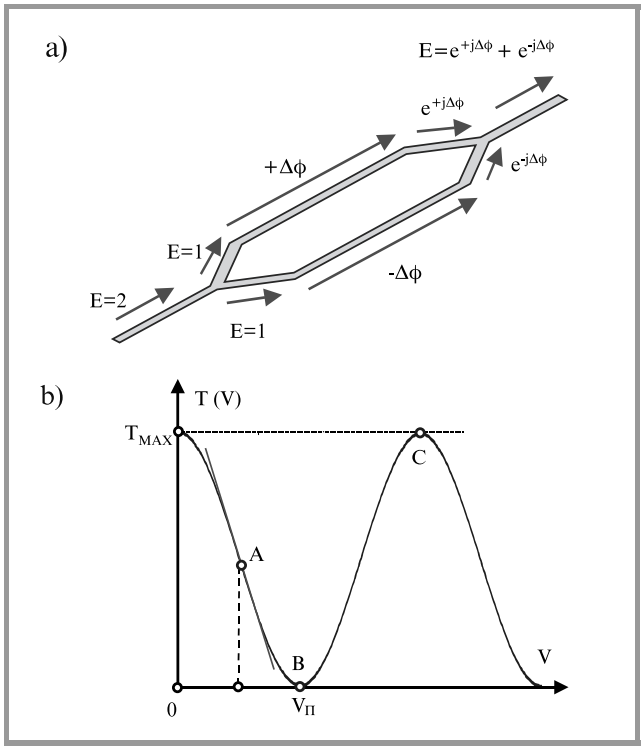


Fig. 2. (a) Illustration of Mach-Zehnder interferometer/modulator operation. (b) Characteristic of power transmission through M-Z modulator.

The transmission can be described as:

$$T(V_{RF}) = \frac{P_{o\,out}}{P_{o\,in}} = \left| \frac{E\left(e^{j\Delta\phi} + e^{-j\Delta\phi}\right)}{2E} \right|^2 = \\ = \cos^2\left[\Delta\phi\left(V_0 + V_{RF}\right)\right] \approx \frac{1}{2} - \Delta\phi\left(V_{RF}\right). \qquad (2)$$

The maximal transmission T_{MAX} of optical power is less than unity because the modulator has its own losses resulting from the planar optical waveguides losses. The formula describing the optical power transmission should include them. When the characteristic value of the modulator voltage (V_π) is considered the transmission can be expressed as (3):

$$T(V) = \frac{P_{o\,out}}{P_{o\,in}} = \frac{T_{MAX}}{2}\left\{1 + \cos\left[\frac{\pi V}{V_\pi}\right]\right\}. \qquad (3)$$

In this expression the voltage $V = V_0 + V_{RF}$ is the superposition of a bias voltage V_0 and a signal voltage V_{RF}. The full shape of the transmission characteristic is shown in Fig. 2b. The power transmission in the point of equality decreases almost to zero. When the symmetry of the power dividing and summing is achieved the ratio $P_{\max}/P_{\min}$ can be obtained up to 1000:1. In the point of inflexion of the $T(V)$ characteristic there is a long straight line section at $V_0 = V_\pi/2$ and with a slope S_{MZ}:

$$S_{MZ} = \left.\frac{\partial T(V)}{\partial V}\right|_{V=V_\pi} = -\frac{\pi T_{MAX}}{2V_\pi}. \qquad (4)$$

The modulator transmission $T(V_{RF})$ at $V_0 = V_\pi/2$:

$$T(V_{RF}) \cong \frac{T_{MAX}}{2} + S_{MZ}V_{RF} = \frac{T_{MAX}}{2}\left(1 - \frac{\pi V_{RF}}{V_\pi}\right). \qquad (5)$$

The pre-bias at point A allows the interferometer to operate as a linear amplitude modulator. This mode of operation is used in the analogue optical links [3]. The voltage V_{RF} can be a sine-form signal or to be a superposition of many carrier frequencies with the modulation side bands.

When the modulator is supplied by a rectangular shape wave with the signal levels corresponding to the points B and C, the device operates as two-state switch. The operation at the bias corresponding to the points B or C has the strongest nonlinearities. This mode of the operation is the most proper in the case of a mixing process.

When the RF signal frequency is being increased, the optical signal transit time τ, through the area of an alternating electric field, becomes comparable with the period T equal to $T = 1/f_{mod}$. When the signal transit time becomes equal to the period T, the modulation effects obtained during one half-period will be disturbed in the next half-period.

Let's assume that:

$$V_{RF} = V_M \sin(\omega t + \Theta). \qquad (6)$$

The presence of the alternating field causes the modulation of the angle $\Delta\phi$:

$$\Delta\phi = m(f)\frac{\pi V_M}{V_\pi}\sin(\omega t + \Theta), \qquad (7)$$

where $m(f)$ – the modulation depth coefficient depends on the modulation signal frequency:

$$m(f) = \left|\frac{\Delta\phi(f)}{\Delta\phi|0|}\right| = m_0\frac{\sin\left(\omega_{mod}\tau/2\right)}{\omega_{mod}\tau/2}. \qquad (8)$$

The finite transit time has the result in the decreasing of the optical signal modulation depth.

Another cause limiting the modulator working bandwidth is the RF signal path mismatching. The RF signal path

is ended with the capacitance which reactance decreases with the increase of the frequency thereby causing the RF path almost short-circuited. A solution for this problem is a travelling wave modulator. In this construction the following situation is obtained. The optical signal speed in the planar waveguide is equal to the modulating microwave signal speed. There are many very interesting solutions for travelling wave modulator structures. However, they will not be described in this paper.

3. Optical-microwave mixing with single M-Z modulator – bias in inflexion point

It was mentioned above that the bias in the inflexion point of the transmission characteristic allows to linearly modulate the optical power transmitted through the modulator. It will be found out that in this point it is possible to work with the optical-microwave frequency mixing process with the use of the nonlinear modulator characteristics [4, 5]. The starting point of the theoretical analysis will be the formula (3), and it will be expressed in a new form (9):

$$T_N(V_0, V_{RF}) = \frac{P_{o\,out}}{P_{o\,in}\,T_{MAX}} = \\ = \frac{1}{2}\left\{1+\cos\left[\phi_0(V_0)+\frac{\pi V_{RF}}{V_\pi}\right]\right\}. \tag{9}$$

In the formula above, $T_N(V_0,V_{RF})$ is the power transmittance of the electro-optical modulator normalized to the maximal power transmittance. By setting appropriate value of the bias voltage to achieve $\phi_0(V_0) = 90^\circ$ and to work in the inflexion point of the transmission characteristic the formula (9) can be written as:

$$T_N(V_{RF}) = \frac{1}{2}\left\{1-\sin\left[\frac{\pi V_{RF}}{V_\pi}\right]\right\}. \tag{10}$$

The optical-microwave mixing process analysed in this point can be performed in the system setup shown in Fig. 3.

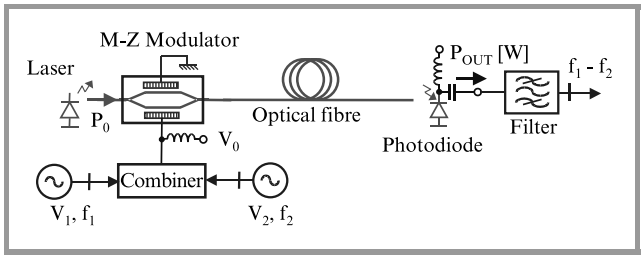


Fig. 3. System to perform optical-microwave mixing process with the use of M-Z modulator.

A combiner and an appropriate bias circuit allow inputting the bias voltage and two alternating sine-form voltages into the modulator. One can write:

$$V_{RF} = V_1 \sin\omega_1 t + V_2 \sin\omega_2 t\,. \tag{11}$$

The normalized transmission T_N can be written as:

$$T_N = \frac{1}{2}\left\{1-\sin\left[X_1\sin\omega_1 t + X_2\sin\omega_2 t\right]\right\}, \tag{12}$$

where X_1 and X_2 are respectively:

$$X_1 = \frac{\pi V_1}{V_\pi}\,,\quad X_2 = \frac{\pi V_2}{V_\pi}\,. \tag{13}$$

With the use of three well-known identity relations, presented in the Appendix, there is possible to expand the transmittance (12) into the Fourier series. This Fourier series expression contains Bessel functions. Considering only the first terms of the expansion one can obtain:

$$T_N = \frac{1}{2} - J_1(X_1)J_0(X_2)\sin\omega_1 t + \\ -J_0(X_1)J_1(X_2)\sin\omega_2 t - J_1(X_1)J_2(X_2)\times \\ \times\left\{\sin\left[(\omega_1-2\omega_2)t\right]+\sin\left[(\omega_1+2\omega_2)t\right]\right\}+ \\ -J_2(X_1)J_1(X_2)\left\{\sin\left[(2\omega_1-\omega_2)t\right]+\right. \\ \left.+\sin\left[(2\omega_1+\omega_2)t\right]\right\}+\ldots\,. \tag{14}$$

The expression above allows analysing and explaining two phenomena related to the nonlinearities of the modulation process.

3.1. Intermodulation distortions during two-tone operation

During the two-tone operation the amplitudes of both sine-form signals are equal to each other $V_1 = V_2$. When the amplitude values are small, then $X_1 = X_2 << 1$. Considering only the first terms of the expansion into the series of the Bessel functions in (14): $J_0(X), J_1(X)$ and $J_2(X)$, according to the formula given in Appendix, one can obtain the following simplified form of the expression (14):

$$T_N = \frac{1}{2} - \frac{X_1}{2}\sin\omega_1 t - \frac{X_2}{2}\sin\omega_2 t + \\ -\frac{X_1X_2^2}{16}\left\{\sin\left[(\omega_1-2\omega_2)t\right]+\right. \\ \left.+\sin\left[(\omega_1+2\omega_2)t\right]\right\}-\frac{X_1^2X_2}{16}\left\{\sin\left[(2\omega_1-\omega_2)t\right]+\right. \\ \left.+\sin\left[(2\omega_1+\omega_2)t\right]\right\}+\ldots\,. \tag{15}$$

The power spectrum of the electrical signal at the photodetector output contains, beside the expected signals at the frequencies ω_1, ω_2 and the amplitudes which are proportional to V_1^2 and V_2^2, also the undesired intermodulation products $(2\omega_1-\omega_2)$ and $(2\omega_2-\omega_1)$ with the amplitudes which are proportional to the third power of the both signal powers at the input of the modulator.

3.2. Products of optical-microwave mixing at bias in the inflexion point

In the optical-microwave mixing process there are also two signals at the frequencies ω_1, ω_2 connected to the modulator. The amplitude of the first of them, called also the signal, is small so the condition $X_1 << 1$ is fulfilled. The second signal at the amplitude V_2 plays role of a heterodyne and usually $V_2 >> V_1$. The process products described above and undesired become now useful products of the frequency mixing. The values of the transmission T_N terms at the frequencies $\omega_1(2\omega_2 - \omega_1)$ are equal to:

$$\begin{aligned} T_{N(\omega_1)} &= -\frac{X_1}{2} J_0(X_2) \sin \omega_1 t \,, \\ T_{N(2\omega_2-\omega_1)} &= -\frac{X_1}{2} J_2(X_2) \sin\left[(\omega_1 - 2\omega_2)t\right] . \end{aligned} \quad (16)$$

The results of the simulation calculations are presented in Fig. 4. It was assumed that the power of the signal at the M-Z modulator input is simply related to the amplitude $V_{1,2}$ and the characteristic impedance Z_0 of the modulator planar transmission line $P_{LO} = V_{1,2}^2/2Z_0$. In calculations it was assumed: $Z_0 = 50\ \Omega$, $V_\pi = 7$ V.

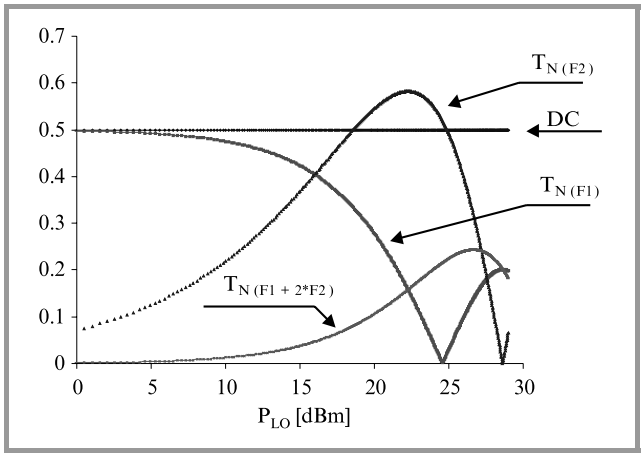


Fig. 4. Products of optical-microwave mixing for circuit shown in Fig. 3 and at bias in the inflexion point. Calculations have been made for $Z_0 = 50\ \Omega$ and $V_\pi = 7$ V.

The analysis of the Bessel functions curves presented in the Appendix allows explaining the role of the heterodyne. The increase of the amplitude V_2 causes decreasing of the amplitude of $T_{N(\omega_1)}$ until it disappears at $X_2 = 2.42$, when $J_0(2.42) = 0$. At the same time the amplitude of $T_{N(2\omega_2-\omega_1)}$ increases obtaining the maximum at $X_{2OPT} = 3.05$, when $J_2(3.05) = 0.486$. The frequency conversion process is the most effective at this value of X_2 and the conversion losses have the smallest value equal about 6 dB.

The power of the signal led to the M-Z modulator is related in a simple way to the amplitude $V_{1,2}$ and the characteristic impedance Z_0 of the planar transmission line of the modulator. The optimal value of the heterodyne power P_{LOOPT}, at which $X_2 = X_{2OPT}$ and the conversion losses are the smallest, can be calculated with the use of the following formula:

$$P_{LOOPT} = \frac{V_2^2}{2Z_0} = \frac{V_\pi^2 X_{2OPT}^2}{2\pi^2 Z_0} \,. \quad (17)$$

One can notice that the lower sideband $(2\omega_2 - \omega_1)$ used in the mixing process corresponds to the harmonic mixing. The upper sideband $(2\omega_2 + \omega_1)$ can be used as well.

4. Optical-microwave mixing with single M-Z modulator – bias in point of maximal transmission

The bias at the maximum of the transmission characteristic, obtained when $\phi_0(V_0) = 0°$, causes also the frequency conversion process to take place. This characteristic is then expressed by:

$$T_N(V_{RF}) = \frac{1}{2}\left\{1 + \cos\left[\frac{\pi V_{RF}}{V_\pi}\right]\right\} . \quad (18)$$

When the RF voltage V_{RF} is a superposition of two parts according to the formula (11), the normalized transmission T_N can be written as:

$$T_N = \frac{1}{2}\left\{1 + \cos\left[X_1 \sin \omega_1 t + X_2 \sin \omega_2 t\right]\right\} . \quad (19)$$

With the use of the relations presented in the Appendix and by limiting analysis to few terms of the expansion, the transmission T_N can be expressed as:

$$\begin{aligned} T_N = \frac{1}{2}\left[1 + J_0(X_1)\, J_0(X_2)\right] + \\ + J_2(X_1)\, J_0(X_2) \cos(2\omega_1 t) + \\ + J_0(X_1)\, J_2(X_2) \cos(2\omega_2 t) + \\ - J_1(X_1)\, J_1(X_2)\left\{\cos\left[(\omega_1 - \omega_2)t\right] + \right. \\ \left. - \cos\left[(\omega_1 + \omega_2)t\right]\right\} + \ldots . \end{aligned} \quad (20)$$

The relationship obtained above allows using interesting applications of this characteristic mode of the operation of the M-Z modulator.

4.1. Operation in frequency doubler configuration

When the modulator has only one signal given at the input

$$V_{RF} = V_1 \sin \omega_1 t \quad (21)$$

then in the power spectrum at the photodetector output the strong product at the frequency $2\omega_1$ is observed, beside the constant product. The amplitude of the term $T_{N(2\omega_1)}$ increases at the beginning proportionally to V_1^2, and reach its maximum at $X_2 = 1.84$, when $J_2(1.84) = 0.581$. At this value of X_2 the frequency multiplying process has the maximal efficiency.

The frequency multiplying process is applied in the fibre-radio systems to overcome the limit resulting from the maximal modulation frequency of the electro-optical modulator.

4.2. Optical-microwave mixing products

The situation is similar like in the Sec. 3.2, in the optical-microwave mixing process two signals at the frequencies ω_1 and ω_2 are connected to the modulator. The amplitude of the first of them, called signal, is small, to fulfil the condition $X_1 << 1$. The second signal at the amplitude V_2, plays role of the heterodyne, and $V_2 >> V_1$. The transmittance term T_N at the frequency according to the lower sideband $(\omega_1 - \omega_2)$ is then expressed as:

$$T_{N(\omega_1-\omega_2)} = -\frac{X_1}{2} J_1(X_2) \sin\left[(\omega_1 - \omega_2)\, t\right] . \qquad (22)$$

The results of the calculations are shown in Fig. 5.

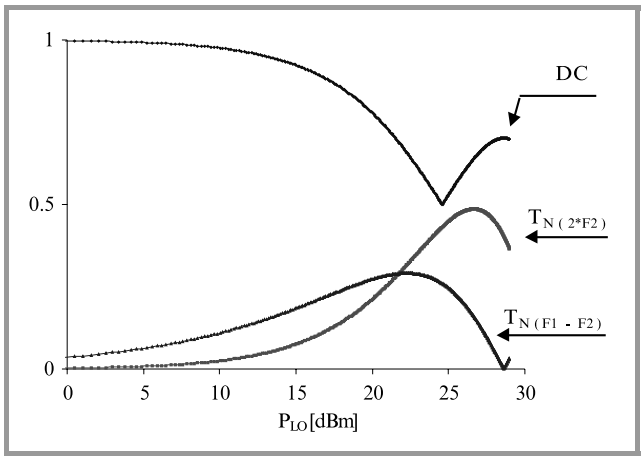


Fig. 5. Products of optical-microwave mixing for circuit shown in Fig. 3 and at bias in point of maximal transmission. Calculations have been made for $Z_0 = 50\ \Omega$ and $V_\pi = 7$ V.

The value of the lower sideband $(\omega_1 - \omega_2)$ amplitude reaches its maximum at $X_2 = 1.84$, when $J_1(1.84) = 0.581$. Because the maximum of function $J_1(X)$ is quite bigger than the maximum of $J_2(X)$, the conversion efficiency in this case is higher than the efficiency described in Sec. 3.2. Also it can be obtained at lower level of the heterodyne power (then at $X_2 = 3.05$, now at $X_2 = 1.84$). The heterodyne power at the optimal value of the conversion losses can be calculated with the use of the formulas $(5 \div 20)$.

5. Optical-microwave mixing with cascade connection of M-Z modulators

Another possibility to obtain the frequency conversion process is the use of the system shown in Fig. 6. Here, the optical signal generated by the laser is transmitted through two M-Z modulators.

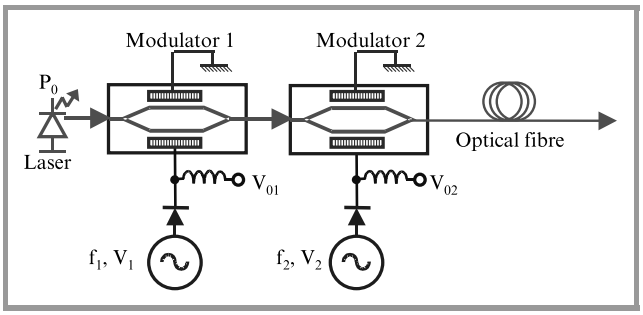


Fig. 6. System to perform optical-microwave mixing process with cascade connection of M-Z modulators

5.1. Products of optical-microwave mixing process- both modulators biased at inflexion points of transmission characteristics

In the case described here the modulators are pre-biased at the voltages V_{01} and V_{02}, which values are set to obtain the operation in the inflexion points of the transmission characteristics. The power transmission of the single modulator is described by the formula (9). Assuming that the modulators are identical, the modulating voltages are equal to: $V_1 \sin \omega_1 t$ and $V_2 \sin \omega_2 t$, respectively, and the values of the variables X_1 and X_2 are defined according to (13). The overall transmission T_{NC} can be expressed as follows:

$$T_{NC} = T_{NI} T_{N2} =$$
$$= \frac{1}{4}\left\{1 - \sin\left[X_1 \sin \omega_1 t\right]\right\}\left\{1 - \sin\left[X_2 \sin \omega_2 t\right]\right\} . \qquad (23)$$

With the use of the identity relations presented in the Appendix there is possible to expand the expression above into the Fourier series. Considering only the most important products at the frequencies ω_1, ω_2, $(\omega_1 - \omega_2)$ and $(\omega_1 + \omega_2)$ one can obtain:

$$T_{NC} = \frac{1}{4} - \frac{1}{2} J_1(X_1) \sin\left[\omega_1 t\right] - \frac{1}{2} J_1(X_2) \sin\left[\omega_2 t\right] +$$
$$- \frac{1}{2} J_1(X_1) J_1(X_2)\left\{\sin\left[(\omega_1 - \omega_2)t\right] + \sin\left[(\omega_1 + \omega_2)t\right]\right\} . \qquad (24)$$

Assuming that the second modulator plays role of the heterodyne, and the first one is driven by the signal small enough to obtain $X_1 << 1$, the transmittance of the product corresponding to the frequency $(\omega_1 - \omega_2)$ can be written as:

$$T_{NC(\omega_1-\omega_2)} = -\frac{X_1}{4} J_1(X_2) \sin\left[(\omega_1 - \omega_2)t\right] . \qquad (25)$$

The results of calculations of some optical-microwave mixing products according to the formula (24) are shown in Fig. 7.

Also in this case the value of the lower sideband $(\omega_1 - \omega_2)$ amplitude reaches its maximum at $X_2 = 1.84$, when $J_1(1.84) = 0.581$, it means that the value of the conversion losses is the smallest and equal to 10.7 dB.

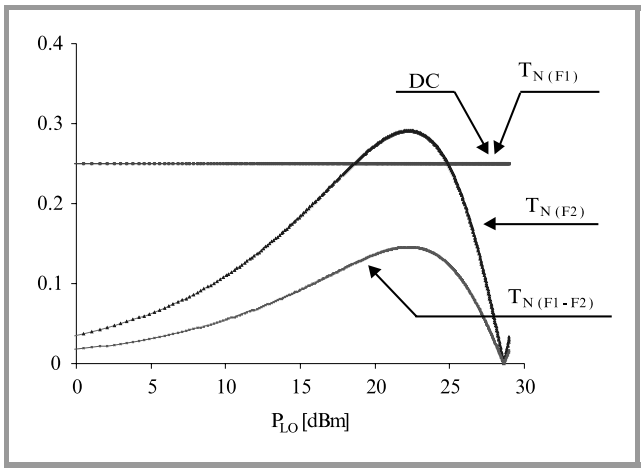


Fig. 7. Products of optical-microwave mixing for circuit shown in Fig. 6. Both modulators are biased at inflexion points. Calculations have been made for $Z_0 = 50\,\Omega$ and $V_\pi = 7$ V.

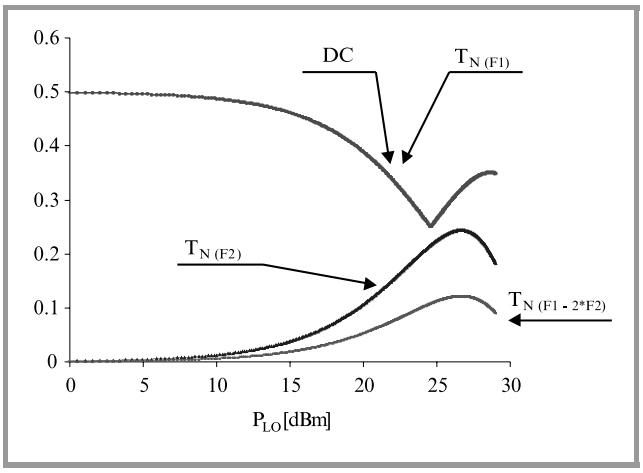


Fig. 8. Products of optical-microwave mixing for circuit shown in Fig. 6. Modulators are biased at different points. Calculations have been made for $Z_0 = 50\,\Omega$ and $V_\pi = 7$ V

5.2. Products of optical-microwave mixing process – modulator biases are different

Very interesting solution is the optical-microwave mixing process performed in the system where the first modulator, called the signal modulator, is biased at the inflexion point of the transmission characteristic, and the second one, the heterodyne modulator, is biased at the point of the maximal transmission. The transmission of the modulators is then described by the formulas (10) and (18). Let's assume that the modulators are identical, and the modulating voltages are: $V_1 \sin \omega_1 t$ and $V_2 \sin \omega_2 t$ respectively, and the values of the variables X_1 and X_2 are defined according to (13). The overall transmission of the cascading connection of the modulators is then equal to:

$$T_{NC} = \frac{1}{2}\left\{1 - \sin\left[X_1 \sin \omega_1 t\right]\right\} \times \times \left\{1 + \cos\left[X_2 \sin \omega_2 t\right]\right\}. \quad (26)$$

The expression above can be expanded into the Fourier series with the use of the formulas given in the Appendix. Considering only the most important terms of the expansion one can obtain:

$$T_{NC} = \tfrac{1}{4}\left[1 + J_0(X_2)\right] - \tfrac{1}{2}J_1(X_1)\left[1 + J_0(X_2)\right]\sin\left[\omega_1 t\right] + + \tfrac{1}{2}J_2(X_2)\cos\left[2\omega_2 t\right] - \tfrac{1}{2}J_1(X_1)J_2(X_2) \times \times \left\{\sin\left[(\omega_1 - 2\omega_2)t\right] + \sin\left[(\omega_1 + 2\omega_2)t\right]\right\}. \quad (27)$$

Assuming $X_1 << 1$, the transmittance related to the product at $(\omega_1 - 2\omega_2)$ can be written as:

$$T_{NC(\omega_1 - 2\omega_2)} = -\frac{X_1}{4}J_2(X_2)\sin\left[(\omega_1 - 2\omega_2)t\right]. \quad (28)$$

The relationship above is almost identical with those obtained in previous cases. The conversion system operates as the harmonic mixer, it can be useful in some cases, the conversion efficiency is a little bit worse, because the maximum of the function $J_2(X_2)$ is smaller, in comparison to the maximum of the function $J_1(X_1)$ (Fig. 8).

6. Conclusions

The analysis presented above shows the possibility of performing the optical-microwave mixing process with the use of the nonlinear abilities of the electro-optical Mach-Zehnder modulators. The values of the microwave heterodyne power level for which the conversion losses reach their minimum are similar to the optimal heterodyne power levels for the quad-diode microwave mixers.
The relationships derived above are helpful and ready-to-use for calculations of the intermodulation distortions concerning the operation in the linear analogue modulation mode.

Appendix

The identity relations used for the calculations:

$$\sin(A + B) = \sin A \cos B + \cos A \sin B, \quad (A1)$$

$$\sin(X \sin\Theta) = 2\sum_{k=0}^{\infty} J_{2k+1}(X)\sin\left[(2k+1)\Theta\right], \quad (A2)$$

$$\cos(X \sin\Theta) = J_0(X) + 2\sum_{k=1}^{\infty} J_{2k}(X)\cos 2k\Theta, \quad (A3)$$

where $J_n(X)$ are Bessel functions.

For the small values of X it is usefully to expand the Bessel function into the series, according to the following formula:

$$J_n(X) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(n+k)!}\left(\frac{X}{2}\right)^{n+2k} \quad n = 0, 1, 2, \ldots . \quad (A4)$$

References

[1] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*. Wiley & Sons, 1991.

[2] K. Noguchi *et al.*, "A broadband Ti:$LiNbO_3$ optical modulator with a ridge structure", *IEEE J. Lightw. Technol.*, vol. 13, no. 6, pp. 1164–1168, 1995.

[3] T. K. Fong *et al.*, "Dynamic range of externally modulated analog optical links", *Proc. SPIE*, vol. 2155–25, pp. 229–240, 1994.

[4] G. K. Gopalakrishnan, W. K. Burns, and C. H. Bulmer, "Microwave-optical mixing in $LiNbO_3$ modulators", *IEEE Trans. MTT*, vol. 41, no. 12, pp. 2383–2391, 1993.

[5] B. Cabon, V. Girod, and G. Maury, "Optical generation of microwave functions", in *Proc. 3rd Int. Summer School OMW'2000*, Autrans, France, 28.08-01.09.2000, pp. 21–32.

Zenon R. Szczepaniak received the M.Sc. degree in optoelectronics from Warsaw University of Technology (WUT), Poland in 1998. Since 1998 he is a Ph.D. student in Microwave Devices Division at the Warsaw University of Technology. His research interests include optically-controlled microwave circuits, modelling of optical-microwave devices and systems for telecommunication.
Institute of Microelectronics and Optoelectronics
Warsaw University of Technology
Koszykowa st 75
00-662 Warsaw, Poland

Bogdan A. Galwas – for biography, see this issue, p. 50.

Paper

Methods for description of the total field propagation in the irregular dielectric waveguides

Ella V. Bekker, Elena A. Romanova, and Marian Marciniak

Abstract — Numerical and semi-analytical methods useful for analysis of the total field propagation in dielectric waveguides with sharp and smooth discontinuities are reviewed. The efficiency of these methods to model the field propagation in such structures is discussed comparatively. As an example, spatial transient regime of the radiation field propagation in planar step-index waveguide excited by a Gaussian beam is simulated numerically by two different beam propagation methods.

Keywords — optical waveguides, numerical modelling, beam propagation methods.

1. Introduction

The principle of operation of many fiber-optic devices is based on the spatial transformation of a guided mode by means of irregularities. If a fundamental mode occurs to be mismatched in the consecutive cross-sections of the guiding structure, spatial transient regime appears. In the smooth structures (tapers), the regime is continuous and depends on the taper adiabaticity. After a step-like discontinuity, radiation field leaks out of the fiber over some definite length and then only the fundamental steady-state mode is left. In the region of the spatial steady-state regime the electromagnetic field near the axis of the waveguide consists mainly of the guided modes. However, in any case the total field distribution is the result of the interference of the guided and the radiation parts.

In order to account for the radiation field and simulate the total field propagation in waveguiding devices, various methods have been elaborated. These are numerous realisations of the beam propagation method (BPM) which has been applied to a variety of waveguide problems. Meanwhile some semi-analytical methods have been proposed based mainly on solving the system of integro-differential equations resulting from the spectral expansion of the Maxwell equation solutions.

The purpose of the present paper is to review and classify the available techniques proved to be useful in the treatment of the total field propagation in irregular guiding structures. The attention will be paid to the single-core structures with smooth or sharp discontinuities of the core radius which are known to be the basic elements of the fiber-optic couplers, junctions, launching devices, etc. As an example, we have considered a planar step-index waveguide which can be treated as a sharp discontinuity when excited by a Gaussian beam. Radiation field propagation was simulated numerically by two different beam propagation methods (FFT-BPM and FD-BPM). The results and comparative analysis are presented in the second part of this paper.

2. Review of the available techniques

2.1. Beam propagation methods

The most commonly used numerical method to solve the scalar wave equation is the split-step Fourier series method which is in fact an extension of the beam propagation method originally developed by Fleck, Morris and Feit [1, 2]. Physically the technique corresponds to replacing the continuous refractive index distribution by an infinitesimally thin lens emerged in a homogeneous reference medium of uniform refractive index.

Conventionally to utilise this method in solution of the wave equation an algorithm is employed which is based on the fast Fourier transformation (FFT). The so-called beam propagation method (FFT-BPM) proved to be an accurate and efficient tool for solving a variety of propagation problems in waveguide geometries involving one or two transverse Cartesian coordinates [3 ÷ 8]. However, in application of this method in cylindrical coordinates one has to cope with the increased storage and reduced efficiency.

A spectral method is described in [2] for solving the paraxial wave equation in cylindrical geometry that is based on expansion of the exponential evolution operator in a Tailor series and use of fast Fourier transforms to evaluate derivatives. The scheme avoids operator splitting altogether and should be useful in the geometries in which the operator splitting may not be possible. It allows the propagation calculation to be performed accurately in cylindrical coordinates on a uniform grid.

The accuracy and applicability of the FFT-BPM have been studied intensively [3, 4] resulting in the conclusion that the FFT-based algorithm becomes critical for high-contrast step-index profiles, i.e. the size of propagation steps must be reduced and the refractive index profile itself must be smoothed out. The main disadvantage of the method is in its structured solution since the field should be written in terms of a coupled set of equations that are solved numerically which require tedious matrix inversions. Further, the effects of nonlinearity and diffraction are assumed to be separated in space whereas in reality both act simultaneously.

An alternate numerical scheme to solve the wave equation is to use a finite difference (FD) approximation [9, 10]. Following the finite-difference beam propagation method (FD-BPM) the wave equation is replaced by a finite-difference scheme. The resulting three-diagonal system of equations is solved by some iterative procedure. FD-BPM has been successfully applied mainly to the analysis of the field propagation in planar waveguides, although nonlinear propagation in a radially symmetric structure was simulated too [11].

The comparative analysis of FFT-BPM and FD-BPM was performed in [9]. FD-BPM occurred to be much more stable with respect to longitudinal step size and number of grid points variations. For comparable accuracy much smaller longitudinal step is necessary to use in FFT-BPM than in FD-BPM, especially in the analysis of the step-index waveguides. Furthermore, the computation time per propagation step for FD-BPM is $4 \div 6$ times less than that of FFT-BPM. However the wide-angle implementation of FD-BPM becomes less accurate if the angle of the structure relative to the propagation direction increases. In spite of the justified feasibility of the FD-BPM for the beam propagation simulations, in the case of the "diagonal dominance" [64] of the tridiagonal matrix coefficients, the algorithm fails. From this point of view the FFT-BPM seems to be more universal technique. Otherwise, there are few types of problems that cannot be analysed accurately using this method. There is a disadvantage consisting of the fact that FFT-BPM requires an uniform transverse grid spacing regardless of how fast are the index and field changes in the transverse direction. This leads to a fast increasing number of grid points for a sufficient calculation accuracy. In the contrary, when using FD scheme, one can choose freely the grid point local density and to increase it at the layer interfaces with high index contrast.

Another problem is the need to apply stair-case approximations to any material boundary of the irregular structure. The boundaries prove to be a non-physical source of numerical noise. The boundary condition used in the conventional BPM is the so-called absorbing boundary condition. The idea is to artificially place a lossy medium at the edges of the computational window to absorb the reflections at the boundary. The major disadvantage of the absorbing boundary condition is that for a specific structure, users have to choose different absorbing parameters. In addition, the computational window size has to be large enough. The transparent boundary condition [7] has distinguished advantages over the absorbing boundary one. First, this condition is automatically set in the numerical algorithm. Therefore, it is not problem dependent. Secondly, a relatively small computational window can be used to increase the efficiency.

Modified finite-difference schemes have been proposed for a dielectric step-like structure in order to stabilise the effect of staircasing on simulated results [12].

Recently the concept of structure related, non-orthogonal coordinate FD-BPM algorithms has been demonstrated to overcome both staircasing and wide-angled propagation problems [$13 \div 20$]. These algorithms adopt the structure related coordinate systems which naturally follow the local geometry of the structure. The algorithm was developed to describe the taper geometry [18] without introducing any approximations. This reduces the number of longitudinal steps that are necessary and the transverse mesh size required.

Early FD mode solvers were based on the scalar wave equation. However, the scalar approximation is accurate only when the refractive index difference is small enough that the system modes are quasi-degenerative. Many different techniques have been adopted for vectorial modelling [$21 \div 23$]. One of the schemes based on the Pade approximation [24, 25] demonstrates considerable improvement in accuracy over the paraxial BPM's.

The semivectorial FD method [26] has been proved to be effective but it does not provide a full vectorial description of the guided field. A novel iterative approach for solving finite-difference equations based on nonlinear iteration have been introduced in [27].

The BPM algorithm breaks down when applied to structures with high refractive index contrast. As an alternative a propagation algorithm based on a finite-difference scheme with a finite element (FE) algorithm is suggested in [28]. This approach is, in general, not restricted to the scalar wave equation for weakly guiding structures or to slow transverse variations in refractive index. The BPM algorithm consists of two separate steps and large step size causes its breakdown. The FE algorithm avoids this problem by not using two separate parts of the propagation. Therefore even large propagation steps are admissible. For the case of a Gaussian beam propagating in a homogeneous medium (for which analytical solutions are available), a much higher accuracy was obtained using the FE method rather than the BPM.

Common FE approaches utilise nodal elements. Nodal elements have been successfully demonstrated for the simulation of slab (two-dimensional) waveguides [29, 30]. Unfortunately, they cannot be applied to three-dimensional waveguides due to the nature of nodal elements as they do not prevent spurious solutions and due to the resulting nonunitary propagation scheme. The additional use of edge elements suppresses non-physical solutions [31]. The combination of edge and nodal elements is called a mixed element approach [32]. A full vectorial simulation of 3D-waveguide structures becomes possible, because spurious modes can be prevented and boundary conditions at interfaces are correctly described. Therefore, mixed elements offer the ability to model waveguides containing lossy media, dielectric and magnetic materials.

In the eigenmode propagation algorithm based on the method of lines (MoL) the eigenvectors of the transverse operator of the wave equation which is discretised by finite differences are used for the evaluation of the propagating field [33]. The method is very accurate, because it directly solves the Helmholtz equation and provides an analytical solution in propagation direction (along the lines). In con-

trast to BPM the field calculations include reflections and radiation from the waveguide discontinuities [34, 35]. Radiation modes are taken into account by using absorbing boundary conditions. In [36] the MoL has been applied for studying the propagation of TE polarised field in a nonlinear planar waveguide with saturable nonlinearity. At low input power there is no practical difference between FD BPM and MoL results. When the power is large enough the comparison shows that the paraxial approximation is not valid.

A new method for studying the propagation of a field through the general waveguiding structure was developed in [37, 38]. It is based on converting the scalar wave equation into a matrix total differential equation using the orthogonal collocation method which may be stated as follows. The partial differential equation is assumed to be satisfied exactly at some points along the radial coordinate. These points are known as the collocation points. The total field is expressed as a linear combination of a set of orthogonal functions. So, by applying the collocation principle, the scalar wave equation can be converted into a set of total differential equations, which can be solved then numerically using any standard procedure, such as the Runge-Kutta method.

2.2. Semi-analytical methods

Numerous techniques exist for analysing irregular optical waveguide systems in which some physical approximations are made since an exact analytical solution is not available. The local modes approach is based on the approximation of an irregular waveguide by the consequence of longitudinally uniform segments [39, 40]. The refractive index profile is assumed to be the one of a regular waveguide defined at the given longitudinal coordinate for every segment. The phase of a local mode is approximated by averaging the longitudinal propagation constant. The local modes have been proved to be suitable only in adiabatic approximation of the field propagation in the waveguides with slowly varying parameters. Since the total field propagating along the irregular waveguide can be expressed as a sum of the discrete set of the guided modes and the integral over the radiation modes, the local modes approach should be generalised on the non-adiabatic case. The total field is to be approximated by the system of the coupled-mode equations for the amplitude of every mode of the field expansion [41 ÷ 43].

The coupled-mode theory was initially used by Marcuse [44] for analysing wave propagation in dielectric waveguide transitions in terms of the guided and radiation modes of the waveguide. It has been used to analyse non-uniform slab waveguides [45] and optical fibers [41]. There are basically two approaches to solve the infinite set of coupled differential equations resulting from this method. On one hand, we can solve a fundamental set of equations numerically by discretising the problem.

Alternatively, we can obtain closed-form expressions using some physical approximations. Three basic approximations are made in coupled-mode theory [46]. First, it is assumed that the waveguide transition is gradual so that, at any point along the length of the transition, the coupling coefficients are small. Second, coupling to reflected guided modes and backward travelling radiation modes is assumed to be negligible. Third, coupling between the different radiation modes is ignored. By making these approximations, one obtains an infinite set of coupled differential equations for the guided mode and radiation-mode amplitudes that need to be solved. By making the assumption that the guided mode amplitude can be treated as a constant, it is possible to integrate directly the radiation-mode amplitudes. However, this approach does not always yield physically meaningful results [46].

To analyse waveguide tapers and step-like structures one can use the step-transition method [46] which was initially used by Marcuse to analyse low-loss tapers [47]. The technique was modified so that it could be used to analyse any taper and step-like structures [46]. It is based on the matching procedure of the transverse fields across the step. A similar mode-matching procedure have been developed that in addition to discrete set of guided modes in subparts of the cross-section includes also continuous spectra of radiation modes [48 ÷ 50]. The method starts by dividing the total cross-section into laterally uniform sections. For each section the complete set of modes is set up. By matching tangential field components at the interface to those of the complete set of modes in the neighbouring section a scattering matrix is constructed which relates the mode amplitudes which are excited in the neighbouring section and the amplitudes of the reflected modes in the section of incidence to the incident mode amplitudes. For the numerical evaluation of the integrals over continuous spectra of radiation modes in this formulation suitable discretisation and normalisation approximates the radiating modes by discrete set of modes.

The general approach to the diffraction problem solution applied to the irregular part of the dielectric waveguide is presented in [51]. It is based on the integral equations formalism [66]. The surface currents are presented as a sum of the uniform part (corresponding to the surface modes of the waveguide) and the non-uniform one. The latter appears due to the irregularities and has a complicated structure. To solve the system of integral equations the Galerkin method was applied. However the realisation of the method demands large computation time. Another technique based on the iterative specifications of the surface mode amplitudes occurred to be more efficient. The method was widely used to calculate and analyse the different types of discontinuities in planar dielectric waveguides [52, 53].

The problem of diffraction of a surface mode by the step-like irregularities of a thin anisotropic waveguide was studied in [54] by the spectral expansion of the total field. The integro-differential equation is transformed onto the Winer-Hopf one and is solved by the method of approximate factorisation [55]. The cases of the mode scattering by a small step in a diameter and by an open end of a semi-infinite

waveguide are analysed in detail. The problem is studied by the variational method in [56] and the both semi-analytical techniques are compared in [57].

In the terms of generalised variational approach a method for obtaining equations for moments of a light beam was proposed in [58]. The method is based on the use of the Galerkin criterion on the basis of a small number of flexible Gaussian modes. The system of ordinary differential equations, with the beam parameters obeying this equation, is easily solved numerically and can be qualitatively analysed [59, 60].

Quasi-analytical method for the analysis of rectangular waveguide structures with step discontinuities is proposed in [61]. It is one of the versions of the Galerkin method in which various weighted polynomials are used as the basis functions. Following this method, one reduces the set of integral equations to the set of linear algebraic equations. The basic point of the technique is evaluation of the matrix elements of the algebraic system, since the accuracy of their computation affects substantially the precision of the final solution. Rapid convergence and high accuracy of the numerical results is guarantied by the method proposed which is based on the separation of the static and dynamic parts of the matrix operator. The method can be applied for the design of different waveguide components: filters, phase shifters, polarisers, etc.

3. The application example

3.1. Comparative analysis of the FD-BPM and the FFT-BPM simulations of the radiation field propagation

Two different beam propagation methods have been applied to simulate the total field propagation in the step-index planar waveguide with infinite cladding excited by the Gaussian beam. One is the FFT-BPM used to solve the Helmholtz equation for the slowly varying amplitude of the total field [7]. The other is FD-BPM [62] applied to solve the paraxial parabolic wave equation, in assumption that the paraxial approximation is valid for a weakly-guiding waveguide. The main goal of the treatment is to evaluate the applicability of the paraxial approximation dealing with guiding structures with sharp contrast of the refraction coefficient. The FD-BPM algorithm is known to fail [64] being applied to solve the Helmholtz equation. Otherwise, the FFT-BPM is free of this disadvantage. On other hand, FFT-BPM is sensitive to refraction index difference in transverse direction. Therefore FFT-BPM breaks down when applied to structures with large index discontinuities.

Propagation of the radiation field excited by a Gaussian beam on an input endface of a single-mode planar waveguide was simulated numerically. The beam was assumed to be parallel to the waveguide layers. The excitation have been considered to be symmetric (beam axis coincides with the input endface center) as well as asymmetric (beam axis is shifted out of the input endface center). The transverse distribution of the incident radiation field was taken as

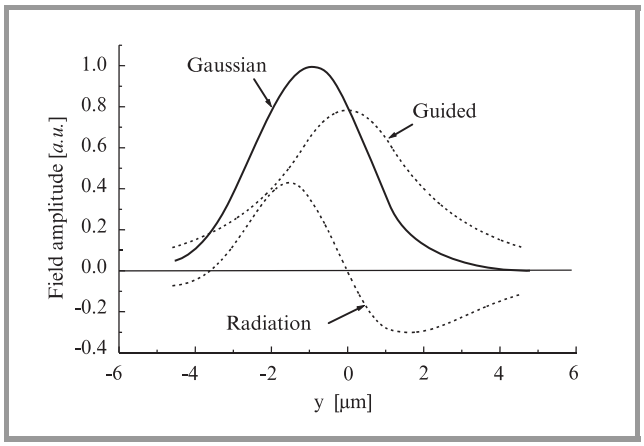


Fig. 1. Input Gaussian beam profile, its modal and radiation parts.

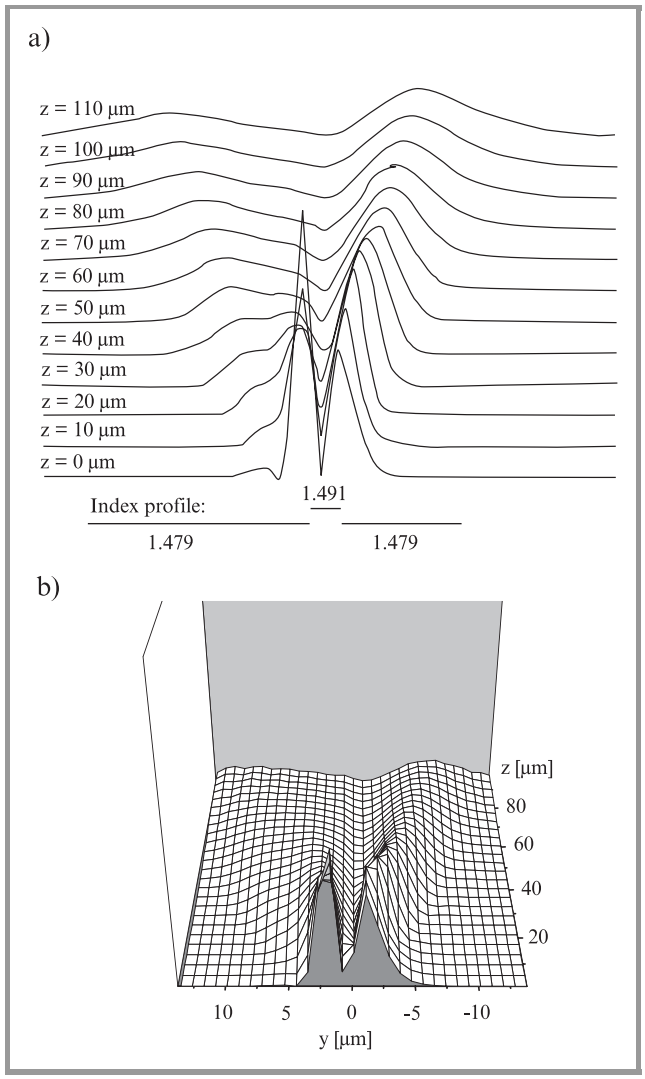


Fig. 2. Spatial transient process of the radiation field propagation simulated by the FFT-BPM (a) and by the FD-BPM (b).

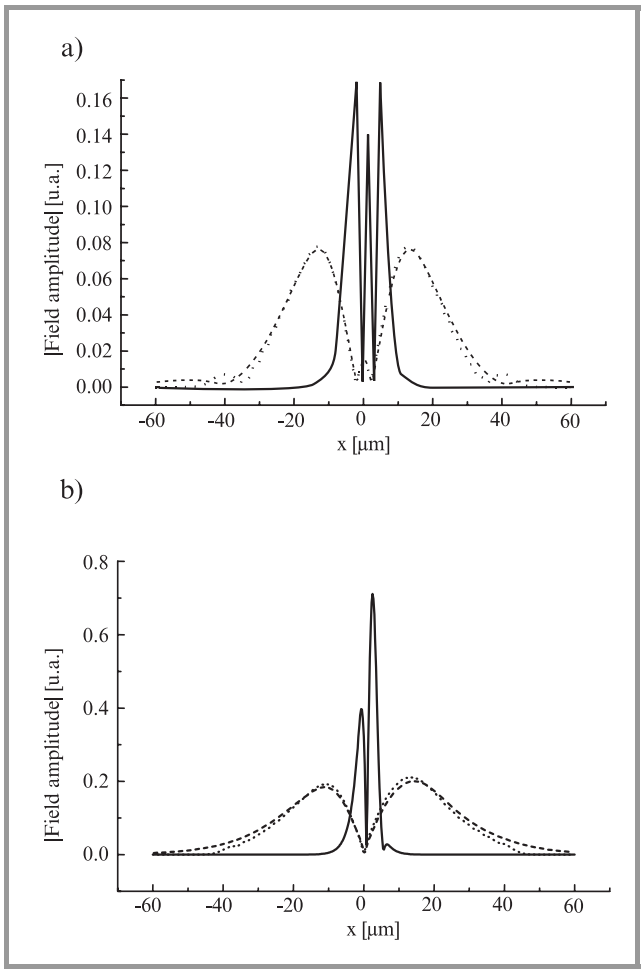


Fig. 3. Transverse distributions of the radiation field at the input endface of the waveguide (solid line), and after the distance of 150 μm, simulated by the FFT-BPM (dashed line) and the FD-BPM (dotted line), respectively. (a) Symmetric excitation (shift = 0), (b) asymmetric excitation (shift = 2 μm).

a difference between the incident field distribution and the modal one [6]:

$$E_{rad} = E_{inc} - E_{\mathrm{mod}}, \tag{1}$$

where E_{mod} is proportional to a fundamental guided mode distribution Ψ_0 normalised to unit power flow, with a complex amplitude α:

$$E_{\mathrm{mod}} = \alpha\Psi_0. \tag{2}$$

The Gaussian beam profile and its modal and radiation parts are shown in Fig. 1.

The refractive indices chosen for the guiding and the cladding layers were $n_1 = 1.491$, $n_2 = 1.479$, respectively. The total width of the guiding layer was $D = 2.3$ μm and the wavelength was $\lambda = 1.53$ μm.

The FFT-BPM computational parameters were chosen in such a way that a part of the non-propagating (evanescent) spectrum and the whole propagating spectrum of the field were included in the k_y-space computational window in order to ensure exact modelling of the field propagation [6].

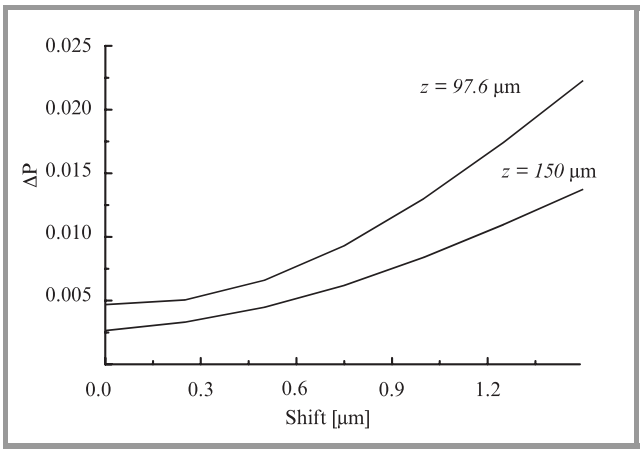


Fig. 4. Normalised power of the radiation field depending on the excitation conditions.

The mesh size was $\Delta x = 0.25$ μm and the longitudinal step size was $\Delta z = 0.5$ μm.

The FD-BPM method has been based on the Crank-Nicolson scheme and an iterative procedure proposed in [63]. The mesh size was $\Delta x = 0.25$ μm and the longitudinal step size was $\Delta z = 2.5$ μm.

Spatial transient process of the radiation field propagation is shown in Fig. 2 simulated by the FFT-BPM (a) and by the FD-BPM (b). The radiation field profiles at the input endface of the waveguide and the ones calculated after a propagation distance of 150 μm are presented in Fig. 3a,b. Additionally, the radiation field power averaged over the region $x = -30$ μm to 30 μm at the distances 96.7 μm and 150 μm was calculated by both techniques. The difference ($\Delta P = (P_{rad}^{FD\text{–}BPM} - P_{rad}^{FFT\text{–}BPM})\, m/P_g$, P_g – guided mode power) is presented in Fig. 4. The results make evidence that the spreading of the paraxial beam (FD-BPM) is less than the wide-angle one (FFT-BPM). The discrepancy

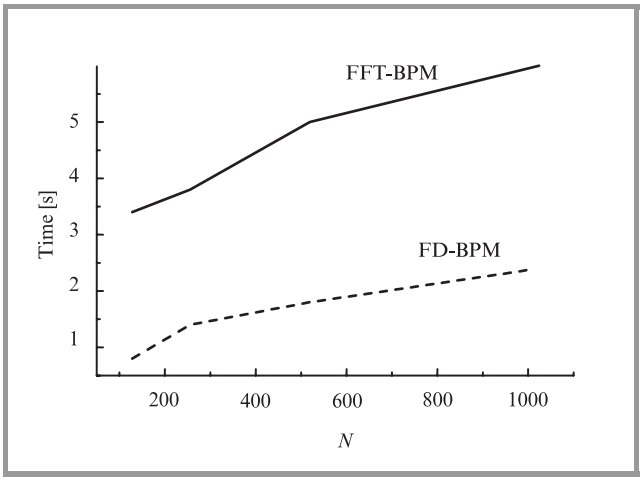


Fig. 5. Calculation time of the numerical simulations of the radiation field propagation (distance = 150 μm, $\Delta z = 0.5$ μm), using the FD-BPM and the FFT-BPM, depending on the number of grid points.

depends on the excitation conditions and grows with the shift of the input beam relative to the input endface center. Additionally, the calculation time required to simulate the radiation field propagation by both these methods was evaluated (Fig. 5). The computation efficiency of FD-BPM simulations is proportional to the number of grid points N while the one of FFT-BPM is of the order NlogN [17].

3.2. Discussion

The modelling of total field propagation by the FFT-BPM and FD-BPM is a clear example that shows some features of these numerical techniques. First, the difference between the results obtained by these methods depend on the radiation field power. If the input beam is strongly mismatched with the modal fields, a great part of the input power excites the radiation field. In this case the results obtained under the paraxial approximation (FD-BPM) differ significantly from the wide-angle ones (FFT-BPM).
Secondly, for comparable accuracy much smaller longitudinal step is to be used in the FFT-BPM than in the FD-BPM. This is one reason for the greater calculation time of the FFT-BPM to achieve the same accuracy. The other reason is that one has to use many FFT expansion terms to describe a complicated spectrum of the radiation field.
As a result, the FD-BPM seems to be more efficient technique comparing with the FFT-BPM being applied to low-contrast quasi-adiabatic structures with small discontinuities.

4. Conclusions

We have reviewed numerical and semi-analytical techniques useful to describe field propagation in all-dielectric guiding structures with smooth and sharp discontinuities. Widely used finite-difference beam propagation method is shown to be efficient tool to deal with weakly-guiding structures with small discontinuities if scalar wave propagation should be traced. Another restriction is that only one-direction field propagation can be simulated. Nevertheless, this method provides good qualitative approach for modelling propagation phenomena in different guiding structures. The review can be used as a reference directory to design fiber-optic elements such as tapers and junctions, as well as for modelling complicated physical phenomena in waveguides (gain, nonlinearity, etc.), when vectorial and wide-angle approaches have to be used.

Acknowledgements

Authors express gratitude for the helpful discussions to Prof. L. A. Melnikov and Dr. S. Sujecki.

References

[1] J. A. Fleck, J. R. Morris, and M. D. Feit, "Time-dependent propagation of high energy laser beams through the atmosphere", *Appl.Phys.*, vol. 10, pp. 129–160, 1976.

[2] M. D. Feit and J. A. Fleck, "Simple spectral method for solving propagation problems in cylindrical geometry with fast Fourier transforms", *Opt. Lett.*, vol. 14, no. 13, pp. 662–664, 1989.

[3] L. Thylen, "The beam propagation method: an analysis of it's applicability", *Opt. Quant. Electron.*, vol. 15, p. 433, 1983.

[4] J. Van Roy, J. van der Donk, and P. E. Lagasse, "Beam propagation method: Analysis and assessment", *J. Opt. Soc. Am.*, vol. 71, p. 803, 1983.

[5] M. Szustakowski and M. Marciniak, "Transmission performance of a 1x2 Ti:$LiNbO_3$ strip waveguide directional coupler", *Opt. Commun.*, vol. 79, pp. 411–415,1990.

[6] M. Marciniak and B. Jaskorzynska, "Radiation field propagation in low-contrast single-mode optical waveguides", *Opt. Quant. Electron.*, vol. 27, pp. 977–985, 1995.

[7] G. R. Hadley, "Transparent boundary conditions for the BPM", *IEEE J. Quant. Electron.*, vol. 28, pp. 363–370, 1992.

[8] D. A. Khalil and S. Tedjini, "Coherent coupling of radiation modes in Mach-Zehnder electrooptic modulators", *IEEE J. Quant. Electron.*, vol. 28, no. 5, pp. 1236–1238, 1992.

[9] Y. Chung and N. Dagli, "An assessment of finite difference beam propagation method", *J. Quant. Electron.*, vol. 26, no. 8, pp. 1335–1339, 1990.

[10] B. M. Andreev, K. P. Eskin, and E. E. Gluhova, "Propagation of radiation along the down-tapered optical waveguide", *Opt. Spectr.*, vol. 61, no. 2, pp. 432–434, 1986 (in Russian).

[11] S. T. Hendow and S. A. Shakir, "Recursive numerical solution for nonlinear wave propagation in fibers and cylindrically symmetric systems", *Appl. Opt.*, vol. 25, no. 11, pp. 1759–1764, 1986.

[12] M. A. Matin, T. M. Benson, P. C. Kendall, and M. S. Stern, "New technique for finite difference analysis of optical waveguide problems", *Int. J. Numer. Model.: Electron. Netw., Dev. Fiel.*, vol. 7, pp. 25–33, 1994.

[13] G. R. Hadley, "Multistep method for wide angle beam propagation", *Opt. Lett.*, vol. 17, pp. 1743–1745, 1992.

[14] Y. Chung and N. Dagli, "A wide angle propagation technique using an explicit finite difference scheme", *IEEE Photon. Technol. Lett.*, vol. 6, pp. 540–542, 1994.

[15] L. Thylen and C. M. Lee, "Beam-propagation method based on matrix diagonalization", *J. Opt. Soc. Am. Ser. A*, vol. 9, pp. 142–146, 1992.

[16] C. Vassallo, "Reformulation for the beam propagation method", *J. Opt. Soc. Am. Ser. A*, vol. 10, pp. 2208–2216, 1993.

[17] H.-P. Nolting and R. Marz, "Results of benchmark tests for different numerical BPM algorithms", *J. Lightw. Technol.*, vol. 13, pp. 216–224, 1995.

[18] P. Sewell, T. M. Benson, P. C. Kendall, and T. Anada, "Tapered beam propagation", *Electron. Lett.*, vol. 32, no. 11, pp. 1025–1026, 1996.

[19] P. Sewell, T. Anada, T. M. Benson, and P. C. Kendall, "Non-standard beam propagation", *Microw. Opt. Technol. Lett.*, vol. 13, no. 1, pp. 24–26, 1996.

[20] P. Sewell, T. M. Benson, T. Anada, and P. C. Kendall, "Bi-oblique propagation analysis of symmetric and asymmetric Y-junctions", *J. Lightw. Technol.*, vol. 15, no. 4, 1997.

[21] L. Xu, W. P. Huang, M. S. Stern, and S. K. Chaudhuri, "Full-vectorial mode calculations by finite-difference method", *Inst. Elect. Eng. Proc. J.*, vol. 141, pp. 281–286, 1994.

[22] W. Huang, C. Xu, S.-T. Chu, and K. Chaudhuri, "The finite-difference vector beam propagation method; analysis and assessment", *J. Lightw. Technol.*, vol. 13, pp. 216–224, 1995.

[23] P. Lusse, P. Stuwe, J. Schule, and H.-G. Unger, "Analysis of vectorial mode fields in optical waveguides by new finite difference method", *J. Lightw. Technol.*, vol. 12, pp. 487–494, 1994.

[24] G. R. Hadley, "Wide-angle beam propagation using Pade approximation operators", *Opt. Lett.*, vol. 17, no. 20, pp. 1426–1428, 1992.

[25] W. P. Huang and C. L. Xu, "A wide-angle vector beam propagation method", *IEEE Photon. Technol. Lett.*, vol. 4, no. 10, pp. 1118–1120, 1992.

[26] M. S. Stern, "Semi-vectorial polarized finite-difference method for optical waveguides with arbitrary index profiles", *Inst. Elect. Eng. Proc. J.*, vol. 135, pp. 56–63, 1988.

[27] S. Sujecki, T. M. Benson, P. Sewell, and P. C. Kendall, "Novel vectorial analysis of optical waveguides", *J. Lightw. Technol.*, vol. 16, no. 7, pp. 1329–1335, 1998.

[28] T. B. Kech, J. B. Davies, and D. Wickramsinghe, "Finite element finite difference propagation algorithm for integrated optical devices", *Electron. Lett.*, vol. 25, no. 3, pp. 514–516, 1989.

[29] Y. Tsuji and M. Koshiba, "A finite element beam propagation method for strongly guiding and longitudinally varying optical waveguides", *J. Lightw. Technol.*, vol. 14, pp. 217–222, Feb. 1996.

[30] Y. Tsuji, M. Koshiba, and T. Tanabe, "A wide-angle beam propagation method based on a finite element scheme", *IEEE Trans. Magnet.*, vol. 33, pp. 1544–1547, 1997.

[31] M. Koshiba and K. Inoue, "Simple and efficient finite-element analysis of microwave and optical waveguides", *IEEE Trans. Microw. Theory Tech.*, vol. 40, pp. 371–377, Feb. 1992.

[32] D. Schulz, C. Glingener, M. Bludszuweit, and E. Voges, "Mixed finite element beam propagation method", *J. Lightw. Technol.*, vol. 16, no. 7, pp. 1336–1341, 1998.

[33] J. Gerdes and R. Pregla, "Beam-propagation algorithm based on the method of lines", *J. Opt. Soc. Am.*, vol. 8, pp. 389–394, 1991.

[34] J. Gerdes, B. Lunitz, D. Benish, and R. Pregla, "Analysis of slab waveguide discontinuities including radiation and absorption effects", *Electron. Lett.*, vol. 28, no. 11, pp. 1013–1015, 1992.

[35] J. J. Gerdes, "Bidirectional eigenmode propagation analysis of optical waveguides based on method of lines", *Electron. Lett.*, vol. 30, no. 7, pp. 550–551, 1994.

[36] M. Bertolotti, P. Masciulli, and C. Sibilia, "MoL numerical analysis of nonlinear planar waveguide", *J. Lightw. Technol.*, vol. 12, no. 5, pp. 784–789, 1994.

[37] A. S. Sharma and S. B. Banerjee, "Method for propagation of total fields or beams through optical waveguides", *Opt. Lett.*, vol. 14, no. 1, pp. 96–98, 1989.

[38] S. Deb and A. Sharma, "Nonlinear pulse propagation through optical fibers: an efficient numerical method", *Opt. Eng.*, vol. 32, no. 4, pp. 695–699, 1993.

[39] A. W. Snyder, "Surface mode coupling along a tapered dielectric rod", *IEEE Trans. Antenn. Propag.*, vol. 13, pp. 821–822, 1965.

[40] J. D. Love, "Spot size, adiabaticity and diffraction in tapered fibers", *Electron. Lett.*, vol. 23, no. 19, pp. 993–994, 1987.

[41] A. W. Snyder, "Coupled-mode theory for optical fibers", *J. Opt. Soc. Am.*, vol. 62, pp. 1267–1277, 1972.

[42] D. Marcuse, "Coupled-mode theory for round optical fibers". *Bell. Syst. Tech. J.*, vol. 52, pp. 817–818, 1973.

[43] V. V. Shevchenko, *Smooth Transitions in Open Waveguides*. Moscow: Nauka, 1969 (in Russian).

[44] D. Marcuse, *Theory of Dielectric Optical Waveguides*. New York: Academic, 1974.

[45] A. Yariv, "Coupled mode theory for guided wave optics", *IEEE J. Quant. Electron.*, QE-9, pp. 919-933, 1973.

[46] P. G. Suchoski Jr. and V. Ramaswamy, "Exact numerical technique for the analysis of step discontinuities and tapers in optical dielectric waveguides", *J. Opt. Soc. Am. A*, vol. 3, no. 2, pp. 194–203, 1986.

[47] D. Marcuse, "Radiation losses of tapered dielectric slab waveguides", *Bell Syst. Tech. J.*, vol. 49, pp. 273–290, 1970.

[48] T. Rozzi, G. Cerri, M. N. Husain, and L. Zappelli, "Variational analysis of the dielectric rib waveguide using the concept of Transition Function and including edge singularities", *IEEE Trans. Microw. Theory Tech.*, vol. 39, pp. 247–256, 1991.

[49] C. J. Smartt, T. M. Benson, and P. C. Kendall, "Free space radiation mode method for the analysis of propagation in optical waveguide devices", *IEE Proc. J.*, vol. 140, pp. 56–61, 1993.

[50] M. Willems and H.-G. Unger, "Analysis and design of optical waveguide components with metallization", in *ICTON'99*, Kielce, Poland, 1999, pp. 41–45.

[51] E. N. Vasiliev, A. V. Polynkin, and V. V. Soloduhov, "Surface wave diffraction at the endface of planar semi-infinite dielectric waveguide", *Radiotech. Electron.*, vol. 25, no. 9, pp. 1862–1872, 1980 (in Russian).

[52] E. N. Vasiliev, A. V. Polynkin, and V. V. Soloduhov, "Excitation of the semi-infinite dielectric plate by the open endface of the planar dielectric waveguide", *Radioelectron.*, vol. 24, no. 2, pp. 60–65, 1981 (in Russian).

[53] E. N. Vasiliev, A. V. Polynkin, and V. V. Soloduhov, "Surface mode scattering from the diameter step of the dielectric waveguides", *Radioelectron.*, vol. 26, no. 2, pp. 72–76, 1983 (in Russian).

[54] A. B. Manenkov, "Propagation of a surface wave along the dielectric waveguide with a jump-like change of parameters". Part 1. "The solution by the method of factorisation", *Radiophys.*, vol. 25, no. 11, p. 1329, 1982 (in Russian)

[55] L. A. Vainshtein, *Theory of Diffraction and the Method of Factorisation*. Moscow: Sov. Radio, 1966 (in Russian).

[56] A. B. Manenkov, "Propagation of a surface wave along the dielectric waveguide with a jump-like change of parameters". Part 2. "The solution by the variational method", *Radiophys.*, vol. 25, no. 12, p. 1484, 1982 (in Russian).

[57] A. B. Manenkov, "A comparison of approximate methods for calculation of wave diffraction at the diameter step in the dielectric waveguide", *Radiophys.*, vol. 28, no. 6, pp. 743–752, 1985 (in Russian).

[58] V. L. Derbov, L. A. Melnikov, and A. D. Novikov, *Kwantovaya Electron.*, vol. 14, no. 12, pp. 2529–2539, 1987 (in Russian).

[59] V. L. Derbov, L. A. Melnikov, A. D. Novikov, and S. K. Potapov, "Transverse pattern formation and spectral characteristics of cw light beams in resonant media: an improved numerical simulation technique and mode analysis", *J. Opt. Soc. Am.*, vol. 7, no. 6, pp. 1079–1086, 1990.

[60] L. A. Melnikov, V. L. Derbov, and A. I. Bychenkov, "Simulation of off-axis Gaussian beams with astigmatism and twisting in a nonlinear waveguide medium", *Opt. Spectr.*, vol. 85, no. 1, pp. 90–94, 1998 (transl.).

[61] V. P. Lyapin, M. B. Manulov, and G. P. Sinyavsky, "Quasi-analytical method for analysis of multisection waveguide structures with step discontinuities", *Radio Sci.*, vol. 31, no. 6, pp. 1761–1772, 1996.

[62] E. A. Romanova, L. A. Melnikov, and E. V. Bekker, "Numerical analysis of the total field propagation in linear and nonlinear single-mode tapered fibres", in *ICTON'99*, Kielce, Poland, 1999, pp. 161–164.

[63] A. N. Tichonov and A. A. Samarsky, *Equations of Mathematical Physics*. Moscow: Nauka, 1972 (in Russian).

[64] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes*. Cambridge, 1986.

[65] R. Mittra, *Computer Techniques for Electromagnetics*. Oxford [etc.]: Pergamon Press, 1973.

Ella V. Bekker – Ph.D. student specialised in laser physics, nonlinear dynamics, nonlinear phenomena in pulse and beam propagation, theory and modelling of optical waveguides. LEOS student member since January 2000.
e-mail: ella@engels.san.ru
Saratov State University
Astrakhanskaya st 83
410026, Saratov, Russia

Elena A. Romanova – Assistant Professor of physics, with 15-years research and teaching experience. Specialysed in theory of optical fibers and modelling of electromagnetic wave propagation in dielectric waveguides.
e-mail:
romanova@optics.sgu.ru
Saratov State University
Astrakhanskaya st 83
410026, Saratov, Russia

Marian Marciniak – for biography, see this issue, p. 55.

INFORMATION FOR AUTHORS

The *Journal of Telecommunications and Information Technology* is published quarterly. It comprises original contributions, both regular papers and letters, dealing with a broad range of topics related to telecommunications and information technology. Items included in the journal report primary and/or experimental research results, which advance the base of scientific and technological knowledge about telecommunications and information technology.

The *Journal* is dedicated to publishing research results which advance the level of current research or add to the understanding of problems related to modulation and signal design, wireless communications, optical communications and photonic systems, speech devices, image and signal processing, transmission systems, network architecture, coding and communication theory, as well as information technology. Suitable research-related manuscripts should hold the potential to advance the technological base of telecommunications and information technology. Tutorial and review papers are published by invitation only.

Papers published by invitation and regular papers should contain up to 15 and 8 printed pages respectively (one printed page corresponds approximately to 3 double-space pages of manuscript, where one page contains approximately 2000 characters).

Manuscript: An original and two copies of the manuscript must be submitted, each completed with all illustrations and tables attached at the end of the papers. Tables and figures have to be numbered consecutively with Arabic numerals. The manuscript must include an abstract limited to approximately 100 words. The abstract should contain four points: statement of the problem, assumptions and methodology, results and conclusion, or discussion, of the importance of the results. The manuscript should be double-spaced on only one side of each A4 sheet (210 × 297 mm). Computer notation such as Fortran, Matlab, Mathematica etc., for formulae, indices, etc., is not acceptable and will result in automatic rejection of the manuscript. The style of references, abbreviations, etc., should follow the standard IEEE format.

References should be marked in the text by Arabic numerals in square brackets and listed at the end of the paper in order of their appearance in the text, including exclusively publications cited inside. The **reference entry** (correctly punctuated according to the following rules and examples) **has to contain**.

From journals and other serial publications: initial(s) and second name(s) of the author(s), full title of publication (transliterated into Latin characters in case it is in Russian, possibly preceded by the title in Russian characters), appropriately abbreviated title of periodical, volume number, first and last page number, year. E.g.:

[1] Y. Namihira, „Relationship between nonlinear effective area and modefield diameter for dispersion shifted fibres", *Electron. Lett.*, vol. 30, no. 3, pp. 262–264, 1994.

From non-periodical, collective publications: as above, but after title – the name(s) of editor(s), title of volume and/or edition number, publisher(s) name(s) and place of edition, inclusive pages of article, year. E.g.:

[2] S. Demri, E. Orłowska, „Informational representability: Abstract models versus concrete models" in *Fuzzy Sets, Logics and Reasoning about Knowledge*, D. Dubois and H. Prade, Eds. Dordrecht: Kluwer, 1999, pp. 301–314.

From books: initial(s) and name(s) of the author(s), place of edition, title, publisher(s), year. E.g.:

[3] C. Kittel, *Introduction to Solid State Physics*, New York: Wiley, 1986.

Figure captions should be started on separate sheet of papers and must be double-spaced.

Illustration: Original illustrations should be submitted. All line drawings should be prepared on white drawing paper in black India ink. Drawings in Corel Draw and Postscript formats are preferred. Colour illustrations are accepted only in exceptional circumstances. Lettering should be large enough to be readily legible when drawing is reduced to two- or one-column width – as much as 4:1 reduction from the original. Photographs should be used sparingly. All photographs must be gloss prints. All materials, including drawings and photographs, should be no larger than 175 × 260 mm.

Page number: Number all pages, including tables and illustrations (which should be grouped at the end), in a single series, with no omitted numbers.

Electronic form: A floppy disk together with the hard copy of the manuscript should be submitted. It is important to ensure that the diskette version and the printed version are identical. The diskette should be labelled with the following information: a) the operating system and word-processing software used, b) in case of UNIX media, the method of extraction (i.e. tar) applied, c) file name(s) related to manuscript. The diskette should be properly packed in order to avoid possible damage during transit.

Among various acceptable word processor formats, TEX and LATEX are preferable. The *Journal's* style file is available to authors.

Galley proofs: Proofs should be returned by authors as soon as possible. In other cases, the article will be proofread against manuscript by the editor and printed without the author's corrections. Remarks to the errata should be provided within two weeks after receiving the offprints.

The copy of the „Journal" shall be provided to each author of papers.

Copyright: Manuscript submitted to this journal may not have been published and will not be simultaneously submitted or published elsewhere. Submitting a manuscript, the authors agree to automatically transfer the copyright for their article to the publisher if and when the article is accepted for publication. The copyright comprises the exclusive rights to reproduce and distribute the article, including reprints and also all translation rights. No part of the present journal may be reproduced in any form nor transmitted or translated into a machine language without permission in written form from the publisher.

Biographies and photographs of authors are printed with each paper. Send a brief professional biography not exceeding 100 words and a gloss photo of each author with the manuscript.

National Institute
of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland

Editorial Office
tel. +48(22) 872 43 88
tel./fax:+48(22) 512 84 00
e-mail: redakcja@itl.waw.pl
http://www.itl.waw.pl/jtit