TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY
Julkaisu 545 • Publication 545

Andrey Krendzel

Network Planning Aspects for 3G/4G Mobile Systems

Tampere 2005

Tampereen teknillinen yliopisto. Julkaisu 545
Tampere University of Technology. Publication 545

Andrey Krendzel

Network Planning Aspects for 3G/4G Mobile Systems

Thesis for the degree of Doctor of Technology to be presented with due permission for public examination and criticism in Tietotalo Building, Auditorium TB109, at Tampere University of Technology, on the 23rd of September 2005, at 12 noon.

Tampereen teknillinen yliopisto - Tampere University of Technology
Tampere 2005

ISBN 952-15-1436-1 (printed)
ISBN 952-15-1828-6 (PDF)
ISSN 1459-2045

ABSTRACT

The thesis “Network Planning Aspects for 3G/4G Mobile Systems” is focused on two topics.

The first part of the research is mainly devoted to *Core Network (CN) planning aspects for the third generation (3G) wireless systems*. This is because of the fact that a large diversity of 3G services based on packet switching leads to dramatic changes in the characteristics and nature of the data traffic. In particular, there will be a considerable increase in the rate of transactions and in the total traffic intensity on Packet Switched Core Network (PS CN) domain elements of the 3G networks, especially, when the Internet Protocol (IP) Multimedia CN Subsystem (IM-subsystem) is involved. Besides, it is necessary to take into account the fact that the multiservice traffic in the 3G networks may have the self-similar nature due to the high variability of burstiness. It may lead to congestion situations and packet-drops in the 3G networks.

As a result of these facts, some 3G network planning problems arise and become more complex than ones relating to the current mobile systems. The following ones are considered in the first part of the thesis: the estimation problem of prospective 3G users, the prediction problem of 3G data traffic characteristics, the problem of performance evaluation of IM-subsystem elements.

It should be emphasized that the CN evolution is quite conservative and it is supposed that the enhanced CN of the 3G systems will be able to support the functionality of services of the fourth generation (4G) mobile systems. It is expected that the major changes in 4G systems will occur in the Radio Access Network (RAN). For this reason, the second part of the research deals with *RAN planning aspects for the 4G wireless systems*.

In particular, the 4G systems should be designed to offer higher bit rate channels (up to 100 Mb/s) and accommodate a significantly larger amount of traffic than the 3G systems. These requirements will make the 4G RAN different from current RANs and will innovate in its architecture. In the 4G systems, a new ring topology of a RAN physical links configuration may be applied in contrast to RANs of current wireless systems. In the second part of the study the RAN ring topology of physical links configuration is estimated from the viewpoint of cost and reliability parameters and the obtained parameters are compared with those of other topologies. As a result, some recommendations on using different topologies of physical links depending on the number of Base Stations (BSs) in the 4G RAN are developed. Besides, the problem of the minimum-cost configuration of physical links between BSs in the 4G RAN ring topology is considered.

PREFACE

The research work for this thesis has been carried out during the years 2003-2005 at the Institute of Communications Engineering of Tampere University of Technology, Finland. This thesis was financially supported by Tampere Graduate School of Information Science and Engineering (TISE), NOKIA Foundation, and Finnish Cultural Foundation. Their support is appreciatively acknowledged.

First I would like to express my sincere gratitude to my supervisor Professor Jarmo Harju for his encouragement, wise guidance, valuable advices, continuous support and infinite tolerance.

I wish to express my sincere appreciation to the reviewers, Professor Heikki Hämmäinen from the Networking Laboratory, Helsinki University of Technology, Finland and Professor Timo Hämäläinen from the Department of Mathematical Information Technology, University of Jyväskylä, Finland for their extremely useful and detailed comments and suggestions on the final version of the manuscript.

I am also truly grateful to Professor Markku Renfors of TISE director and Dr. Pertti Koivisto of TISE coordinator for giving an opportunity for me to be involved in a major program to advance doctoral studies in the Tampere region.

Special thanks to Professor Sergey Lopatin and Dr. Vera Derzhavina from St. Petersburg Research and Development Institute of Telecommunications (LONIIS), Russia for fruitful, long-term co-operation and co-authoring of my publications.

I wish to thank my colleagues Roman Dunaitsev, Yevgeni Koucheryavy, Dmitri Moltchanov in Finland for interesting scientific and nonscientific tea-table debates.

I am most grateful to Tarja Erälaukko, Sari Kinnari, Elina Orava for taking the time and efforts to settle practical matters.

Finally I take this opportunity to express my deepest gratitude to my wife Marina, my son Andrey, my mother Lubov and my father Valery for their love, kindness, support, inspiration, tolerance and understanding during the preparation of this thesis.

Tampere, August 2005

Andrey Krendzel

CONTENTS

LIST OF PUBLICATIONS

P1. Krendzel A., Koucheryavy Y., Derzhavina V., Harju J. "*Method for Estimating the Number of Potential 3G Users*", in proceedings of 10-th Open European Summer School on the Advances in Fixed and Mobile Networks (EUNICE – 2004), Tampere, Finland, June 14-16, 2004.

P2. Krendzel A., Koucheryavy Y., Lopatin S., Harju J. "*Method for Estimating Parameters of 3G Data Traffic*", in proceedings of the IEEE International Conference on Communications (ICC-2004), Paris, France, June 20 – 24, 2004.

P3. Krendzel A., Derzhavina V., Lopatin S. "*Method for Estimating Parameters of NGN traffic*", in proceedings of the International Conference on Next Generation Teletraffic and Wired/Wireless Advanced Networking (NEW2AN), pp. 50-54, St.-Petersburg, Russia, February 02-06, 2004.

P4. Koucheryavy Y., Krendzel A., Lopatin S., Harju J. "*Performance Estimation of UMTS Release 5 IM-Subsystem Elements*", in proceedings of the 4-th IEEE Conference on Mobile and Wireless Communications Network (MWCN 2002), Stockholm, Sweden, September 9-11, 2002.

P5. Koucheryavy Y., Krendzel A. "*Analysis of Web Traffic and Users' Behavior Modeling During Busy Hour*", in proceedings of the 6-th Open European Summer School (EUNICE – 2000), University of Twente, Enschede, the Netherlands, September 13-15, 2000.

P6. Krendzel A., Koucheryavy Y., Harju J. "*Cost and Reliability Estimation of Radio Access Network Structures for the 4G Systems*", in proceedings of the IEEE 58-th Vehicular Technology Conference (VTC 2003-Fall), the Symposium on Wireless Communications: 3G and Beyond, Orlando, USA, October 6-9, 2003.

P7. Krendzel A., Koucheryavy Y., Harju J. "*Radio Access Network Topology Planning for the 4G Networks*", in proceedings of the 5th European Wireless Conference Mobile and Wireless Systems beyond 3G (EW-2004), Barcelona, Spain, February 24 – 27, 2004.

P8. Krendzel A., Koucheryavy Y. "*Remote Units Location Problem in Multiservice Access Network*", in proceedings of International Conference on Telecommunications (ICT-2002), Beijing, China, pp. 496-500, June 23-26, 2002.

ABBREVIATIONS

2G – Second Generation
3G - Third Generation
3GPP - Third Generation Partnership Project
4G – Fourth Generation
ATM - Asynchronous Transfer Mode
BH – Busy Hour
BS – Base Station
BSC – Base Station Controller
BSS - Base Station System
BTS – Base Transceiver Station
CAMEL – Customized Application for Mobile Network Enhanced Logic
CDMA 2000 – A CDMA System in North America
CN – Core Network
CS - Circuit Switched
CS CN - Circuit Switched Core Network domain
CSPDN – Circuit Switched Public Data Network
EDGE – Enhanced Data Rates for GSM Evolution
ETSI - European Telecommunication Standard Institute
GGSN – Gateway GPRS Support Node
GDP – Gross Domestic Product
GERAN – GSM/EDGE Radio Access Network
GPRS – General Packet Radio Service
GSM - Global System for Mobile communications
HSS – Home Subscriber Server
IM-subsystem (IM CN) – IP Multimedia Core Network Subsystem
IMT-2000 - International Mobile Telecommunications System – 2000
IN - Intelligent Network
IP – Internet Protocol
IPv4 – Internet Protocol Version 4
IPv6 – Internet Protocol Version 6
IS-95 – North American Version of the CDMA Standard
ISDN – Integrated Services Digital Network
ITU - International Telecommunication Union
I_u – UMTS Interface Between 3G-MSC/SGSN and RNC
MAP – Mobile Application Part
MM – Mobile Management
MS – Mobile Station
MSC – Mobile Switched Center
N-ISDN – Narrowband ISDN
NGN – Next Generation Network
NMS – Network Management Subsystem
NPV – Net Present Value
NSS – Network Sub System
PDC – Pacific Digital Communication
PDSN – Packet Data Serving Node
PLMN - Public Land Mobile Network
PS - Packet Switched
PSCN - Packet Switched Core Network domain
PSTN - Public Switched Telephone Network

QoS – Quality of Service
RAN – Radio Access Network
RNC – Radio Network Controller
RNS – Radio Network Subsystem
SGSN – Serving GPRS Support Node
SIP – Session Initiation Protocol
SS7 – Signaling System No. 7
TDMA – Time Division Multiple Access
UE – User Equipment
UMTS - Universal Mobile Telecommunication System
U_m – Radio Interface for GSM BSS
UTRAN – Universal Terrestrial Radio Access Network
U_u – Radio Interface for UTRA
VAS – Value Added Service Platform
VHE – Virtual Home Environment
VLR - Visit Location Register
VoIP – Voice over IP
WAP – Wireless Application Protocol
WCDMA – Wideband Code Division Multiple Access

List of Tables

List of Figures

1 INTRODUCTION

1.1 Scope and the objective of the research

There has been an evolution in wireless communications almost every ten years. The first generation (1G) in 1980s and the second generation (2G) mobile systems in 1990s have been oriented mainly for providing circuit-switched (CS) services to users. The 2G subscribers have used the rate for data transfer up to 14 kb/s as a maximum. In 1996, European Telecommunications Standards Institute (ETSI) decided to enhance 2G GSM standard in annual Phase 2+ releases that incorporate the third generation (3G) features such as General Packet Radio Service (GPRS) and Enhanced Data Rates for GSM Evolution (EDGE). The data rates for users of the systems are limited to 384 kb/s [1].

Universal Mobile Telecommunications System (UMTS) as the 3G mobile system will be introduced during first decade of new century. It is specified by ETSI and the world-wide 3G Partnership Project (3GPP) within the framework defined by the International Telecommunication Union (ITU) and known as International Mobile Telecommunications-2000 (IMT-2000). The 3G systems should offer the data rates up to 2 Mb/s depending on mobility/velocity.

ETSI and 3GPP are introducing UMTS in phases and annual releases. UMTS Rel'3 (sometimes called as Rel'99) is a 3G GSM successor standard using the GSM Phase 2+ enhanced core network (CN). The most important evolutionary step toward UMTS is to introduce a packet switched core network (PS CN) domain. The main function of the PS CN domain is to support all services (GPRS, WAP, etc.) provided to both GSM subscribers and UMTS users [1].

The following phases after Rel'3 specify how voice and multimedia can be supported by IP technology. It is characterized by creating of the Internet Protocol (IP) Multimedia Core Network Subsystem (IM-subsystem), which comprises all PS CN domain elements for providing telecommunication services within UMTS Rel'4,5,6. The IM-subsystem contains a uniform way to maintain Voice over IP (VoIP) calls and offers a platform to multimedia services. The examples of IM services are voice telephony, real-time interactive games, videotelephony, instant messaging, emergency calls, multimedia conferencing [2]. In the UMTS Rel'5,6 the traffic coming from Radio Access Network (RAN) to the CN is supposed to be always packet switched; and here the question is whether it is real-time or non-real-time [1].

The next step of wireless communications evolution is the fourth generation (4G) of mobile communication systems (the systems beyond IMT-2000). Now it is difficult to predict when the 3G evolution ends and the 4G really starts [1]. The 4G systems should offer significantly higher bit rate than 2 Mb/s, have high capacity with a low bit cost and be able to support all types of telecommunication services from the viewpoint of multimedia communications [3]. It is supposed that on the CN side of the 4G systems the main purpose is to minimize changes and utilize the 3G CN elements and the 3G CN functionality as much as possible [1]. The CN development is summarized in the Table I.

TABLE I. CORE NETWORK DEVELOPMENT

WIRELESS SYSTEMS	CORE NETWORK DOMAINS
2G	CS CN
2G phase 2 +	CS CN and PS CN
3G (UMTS Rel'3)	CS CN and PS CN (enhanced 2G phase 2+ CN)
3G (UMTS Rel'4)	CS CN, PS CN, IM CN
3G (UMTS Rel'5,6)	IM CN
4G	IM CN (enhanced 3G CN)

There are some important features of the global evolution process in wireless communications.

The 3G wireless systems should be designed to support for a high-speed transfer of a large amount of multimedia information between users. One of the main properties of the data traffic in the 3G systems is a large diversity depending on the profile of services provided to 3G users. It is expected that the traffic in the 3G systems will expand considerable [4]. The growing data/multimedia traffic leads to increasing the total load on network subsystem elements. Moreover, traffic patterns generated by 3G services may be quite different from traditional Poisson models used for circuit switched voice traffic. When modeling packet-switched multiservice networks it is necessary to take into account the notion of self-similarity [5-8]. Due to the high variability of burstiness of the traffic, the use of the classical teletraffic theory for performance evaluation of PS CN domain elements may give essential faults; in particular, the network parameters can be underestimated. Such faults are unacceptable when IM-subsystem planning as well, therefore, principles of the teletraffic theory cannot be applied in this case.

Due to above reasons, the following 3G network planning problems occur:

- the prediction problem of a demand for 3G services;
- the estimation problem of 3G data traffic parameters;
- the problem of performance evaluation of IM-subsystem nodes taking into account the self-similar nature of the multiservice traffic.

It is worth mentioning that the efficient approaches to solve these problems taking into consideration the above-mentioned aspects of 3G wireless systems have not been presented in the publications concerning 3G CN network planning [9-12].

One more feature of the evolution in wireless communications concerns changes in the 4G wireless systems. It is seen from the Table I that the CN evolution is quite temperate. From the viewpoint of functional capabilities the enhanced CN of the 3G systems will be able to support 4G services [1]. So, it is expected that the RAN will undergo the main changes, in particular, on a physical transmission layer. In the 4G systems, because there will be a need to deal with the enormous amount of traffic, the Base Station (BS) cell size is supposed to be smaller than that of 3G systems [13]. Therefore, the 4G RAN will comprise more BSs and more frequent handover will occur resulting in a heavy load on the links between such elements of RAN as Remote Network Controller (RNC) and BSs, suggesting changes to the RAN architecture [3,13,14].

So, for the 4G RAN a new and innovative topology of physical links configuration has been proposed in addition to traditional topologies using in RANs of current wireless systems [3,13,14]. The topology is called the “cluster-cellular” or, in other words, it is known as the ring topology. In such topology BSs are grouped in a “cluster” and there is a “cluster-main” BS connected to the RNC. BSs in a cluster may be connected to each other by optical fiber links that are preferred as the dominant links to construct the 4G RAN from the viewpoint of link capacity [15]. Note that the 4G RAN ring topology has been proposed and analytically argued with the viewpoint of a load and routing capabilities only [3,13,14]. It is obvious that it is very important to analyze the ring topology of the RAN physical links with respect to its reliability and cost as well, to compare it with other topologies and give recommendations on applying of different ones in the 4G RAN. Besides, it is worthwhile to determine the optimal configuration of physical links between BSs in the ring topology to decrease the 4G RAN deployment cost.

Taking into account the considerations above, the research is concentrated on two subjects. The first part of the research is focused mainly on the 3G CN planning problems. The second part of the research deals with the 4G RAN topological planning problems.

The main objective of the research is to develop solution methods for the above-mentioned problems. It will enable planning 3G/4G networks in such a way that both technical and economical advantages can be achieved when constructing and exploiting the networks.

1.2 Outline of the thesis

We start the first part of the thesis devoted to 3G CN planning with the estimation problem of the number of prospective 3G users. The problem may be considered as the initial one for network planning. The successful solution of the problem gives a basis to solve other topical 3G network planning problems.

After that, the prediction problem of the main parameters of data traffic generated by 3G users is considered. In particular, the method for estimating parameters of the 3G user traffic on a level of PS CN domain is developed. In accordance with the parameters obtained as a result of the problem solution it is possible to adjust performance measures of 3G network equipment in order to ensure Quality of Service (QoS).

In the end of the first part of the research the evaluation problem of performance measures of 3G network equipment is considered. So, the method for performance evaluation of IM-subsystem nodes in UMTS Rel’ 5 is developed.

The main purpose of second part of research devoted to 4G RAN planning is to work out the methods that allow reducing the 4G RAN deployment cost taking into account the reliability requirements.

The first problem deals with estimating reliability and cost parameters of different 4G RAN topologies of physical link configuration. In particular, the parameters of one ring, multi-ring and radial topologies are considered and compared with each other.

Then, the configuration problem of physical links between such 4G RAN elements as BSs with a minimal construction cost for the ring topology is considered. The problem arises if the ring topology is selected for 4G RAN planning as a result of the previous problem solution.

Thus, there are the main five interdependent problems that are considered in the thesis. In next subsection we discuss about research methodology for solving these problems.

1.3 Research methodology

The principles of teletraffic theory, probability theory, mathematical analysis, graph theory, economics are used in the thesis to achieve the appropriate results.

So, the estimation problem of the number of potential 3G users requires the application of principles from several disciplines such as telecommunications, mathematics and economics taking into account that a demand for new telecommunication services depends on both Gross Domestic Product (GDP) and its distribution within society. In particular, the method for estimating the number of prospective 3G users is based on the application of the Pareto law, the Lorenz curves, the Jipp curve and the Gini coefficient [16-22].

Since one of the main characteristic futures of 3G wireless systems is to provide a large diversity of services to users it is reasonable for the prediction of the main data traffic parameters to make decomposition of 3G services into some subsets in accordance with average amount of transferred data per transaction and distribution of 3G users into some subgroups in accordance with their demand for 3G services from different subsets of services. Inequality of the distribution of 3G users is taken into account by different values of the Gini coefficient for each subset of 3G services. Such approach allows segregating several segments from the common flow of transactions initiated in busy hour by support procedures of 3G services. After that, the main parameters of data traffic are determined for each segment of data traffic and the total traffic intensity on the PS CN domain is estimated.

The method for performance evaluation of IM-subsystem nodes in 3G wireless systems should take into consideration the influence of self-similar input. For this reason, it is proposed to analyze the performance measures of IM-subsystem nodes on a basis of FBM/D/1/W queueing system, where FBM is Fractional Brownian Motion [23-25].

Topological configuration of 4G RAN physical links is one of the most important problems of network planning because it will determine the long-term performance and service quality of 4G systems. Many factors influence on reliability and cost of RAN network topologies. It is proposed to make a quantitative estimation of reliability and cost parameters of different 4G RAN topologies including the new ring topology on a basis of probability theory principles.

The optimization problem of physical link configuration between BSs in 4G RAN presents an NP-complete problem. The configuration problem for the new ring topology of 4G RAN is to find the shortest path (from the cost viewpoint) beginning and ending in the same node corresponding to the “main” BS location. The dynamic programming is used to provide an acceptable solution for the problem since such approach has some advantages in comparison with other algorithms for the shortest path searching (see chapter 6).

2 ESTIMATING THE NUMBER OF PROSPECTIVE 3G USERS

The prediction problem of a demand for 3G services in a region before deployment of 3G network equipment may be considered as the first one for 3G network planning. Information about the demand for 3G services is very useful for network operators, service providers and network equipment manufacturers. Besides, the solution of the problem enables forming initial data for other 3G network planning problems. In this chapter the method for the preliminary estimation of the potential number of 3G users in a region is considered. It is based on relationships between the Pareto law, Gross Domestic Product (GDP), income distribution within population, and infocommunication density. *The method is presented in detail in* [P1].

2.1 Approach for estimating the number of prosperous 2G users using the Pareto law

The Pareto law was established on a basis of empirical data concerning income distribution in different countries [16,17]. This law has a mathematical formulation and shows a relationship between each income level and the number of people who receive more than that income. Note that the Pareto law is applied not only for economics. It is just for other fields of human activity that are characterized by statistical nature and have distributions, which are nonconvergent to the normal (Gaussian) law [26]. In particular, the Pareto law may be extended for estimating a demand for various services of infocommunication infrastructure, including telecommunication services. This is because of the fact that the demand for the services depends on both GDP and its distribution within society and there is the relationship between a demand for telecommunication services, labour productivity, distribution of incomes between individuals, and GDP [27]. It is proved in [27] that the relationship between a demand for telecommunication services and income distribution is close to the Pareto law.

In [16] Pareto introduced income distribution in the following view

$$R(x) = x^{-\alpha}, \quad 1 < \alpha < \infty, \tag{2.1}$$

where $R(x)$ is the number of individuals who have income more than x, α is the distribution parameter called the Pareto parameter.

The random variable x may be presented as the normalized income that equals to the ratio G/G_{min}, where G is one of income values, G_{min} is the minimum income value.

In accordance with (2.1) the higher the individual income level in a subgroup the less individuals are in the subgroup. In terms of integral distribution it is possible to define (see [P1]) the integral function of probability distribution $F(x)$ and the probability density $w(x)$ with integral distribution $F(x)$. Then, it is possible to get the n-th moment of distribution of the random variable given by the probability density $w(x)$, $M_n = \alpha / (\alpha - n)$. When $\alpha > 1$ the n-th moment of distribution does not exist and the average value of random variable distributed by the Pareto law is defined only. Therefore, the Pareto distribution belongs to the type of Cauchy distributions and is nonconvergent to the normal one [26].

Using the normalization rule the minimum income value may be expressed as $G_{min}(\alpha) = G_0/M(\alpha)$, where G_0 is GDP per capita, $M(\alpha) = \alpha/(\alpha - 1)$ is the average value of the income. Thus, the expression (2.1) for estimating the number of individuals who have income more than $x = G/G_{min}$ takes the following view

$$R(G/G_{\min}) = \left(\frac{G\alpha}{G_0(\alpha-1)}\right)^{-\alpha}, \tag{2.2}$$

It is clear that telecommunications as a part of society infrastructure influence on development of society and determine the economical level of society. At the same time, country economics determine the level of telecommunication development. It is known that there is the relationship between infocommunication density and a level of GDP per capita that may be described by the well-known Jipp curve [28]. For example, there is the linear dependence (in the first approximation) between the telephone density and GDP per capita. It is argued in [27] that the similar linear dependence takes place between amount of produced information, generated by society, per an individual and GDP per head. Generally, it is possible to apply the dependence for other characteristics relating to telecommunications and information services. In our case, it is assumed that there is the linear dependence between the penetration level of 2G services and GDP per capita.

Mathematically it may be presented as

$$\Delta = A \cdot G_0^{\,a}, \tag{2.3}$$

where $\Delta = N_{mt}/100$ is the penetration level of 2G services, N_{mt} is the average number of 2G mobile terminals per 100 individuals, A is the normalizing dimension factor, a is the power index. The index a is close to 1, i.e. $a = 1 + a$, $a \ll 1$. So, in first approximation it may be rewritten as the linear dependence $\Delta = AG_0$.

To analyze a demand for telecommunication services it is quite enough to have information about distribution of large incomes [27]. Besides, the area of small incomes has very small influence on statistical characteristics of income distribution in accordance with the Pareto distribution. Note that the expression (2.1) is true for large income distribution and may be interpreted as probability that the income is more than x. Since parameters Δ and G_0 are the average values obtained by averaging many initial and random data, in general case, the following expression is fair $T = AG$, where T may be interpreted as the penetration level of 2G services in the subgroup of the prosperous users.

Thus, it is possible to get the follows

$$k = \frac{T}{\Delta} = \frac{G}{G_0},\ k \geq 1. \tag{2.4}$$

Here, the coefficient k determines the excess of the penetration level of 2G services in the subgroup of the prosperous users above the average value of the 2G penetration level relating to all individuals.

It gives an opportunity to form the expression for determination of the relative number of individuals who have the penetration level of 2G services more or equal than the parameter Δ, or, in other words, we get the expression for estimating the relative number of the prosperous users of 2G services

$$R(k,\alpha) = \left(\frac{k\alpha}{\alpha-1}\right)^{-\alpha}. \tag{2.5}$$

As a rule, the penetration level of 2G services in a region Δ is known and may be found in statistical literature. Therefore, for estimation of the relative number of the prosperous 2G users (when given T) it is necessary to find the Pareto parameter α only.

2.2 The approach for estimating the Pareto parameter

Generally the parameter value α depends on the inequality of income distribution between individuals. With reference to our case, this parameter depends on the inequality of distribution of the number of mobile phones between individuals. Usually the inequality of income distribution in a subgroup of individuals is illustrated by the Lorenz curves [19,22,29,30]. If the Pareto law describes the income inequality in mathematical view then the Lorenz curves show the income inequality in the integral form. The set of the Lorenz curves when $\alpha = 1.05$; 1.16; 1.50; 3.00 is presented in Figure 1. If all individuals have the same income, the Lorenz curve is a straight diagonal line, called the line of equality. If there is any inequality in income then the Lorenz curve falls below the line of equality. In Figure 1 the Lorenz curves show relationships between the normalized (to the parameter value M) current average income value in a subgroup of population Q and the normalized number of individuals in the subgroup $F = 1 - R$.

The derivation of the analytical function $Q(F)$ that allows assigning a set of the Lorenz curves is given in [P1]. The function has the following view [26,31,32]

$$Q(\alpha, x) = 1 - (1 - F(x))^{\frac{\alpha - 1}{\alpha}} . \tag{2.6}$$

The set of the Lorenz curves shows that with increasing the parameter value α income in a subgroup is becoming more evenly distributed, i.e. the more α the closer a Lorenz curve to the line of equality. It is worth mentioning that the relationships $Q(\alpha, x)$ are the generalization of the well-known inequality of income distribution between population that can be presented as $\rho/(1-\rho)$, $\rho < 0.5$. The ratio means that in many aspects of human activities the small part of population (ρ %) owns the most part of cumulative income $(1 - \rho)$ %. The famous Pareto rule 20/80 (20 percent of the people owned 80 percent of the wealth) corresponds to the Pareto law given by the Lorenz curve $Q(\alpha,x)$ with $\alpha = 1.16$ ($\rho = 0.2$, $M = 7.25$). The value $\alpha = 1.5$ is usually assigned as the mean statistical estimation. It corresponds approximately to the parameter of the inequality $\rho = 0.32$ and the rule 30/70 [26].

As a rule the Gini coefficient is considered in statistical literature as a measure of the income inequality [26]. The Gini coefficient is the integral parameter characterizing deviation of a Lorenz curve from the equal income distribution line. The coefficient is numerically equal to the doubled area enclosed by the line of equality $Q = F$ and the Lorenz curve $Q(F)$ [26].

It should be emphasized that in each stage of development of 2G systems a number of 2G subscribers is known only on a basis of statistical information. In other words, the number of individuals in a subgroup that has 100% even coverage of 2G services, i.e. has completely even distribution of 2G services, is known. It is obvious that the Lorenz curve in this case is transformed into the broken line (on the assumption that one subscriber has only one mobile terminal). It may be expressed as follows

$$Q(\Delta, F) = \begin{cases} 0, & 0 < F < 1 - \Delta \\ \dfrac{F - 1 + \Delta}{\Delta}, & 1 - \Delta \le F \le 1 \end{cases}, \tag{2.7}$$

The view of the distribution (2.7) when the penetration level $\Delta = 0.3$ is shown in Figure 1.

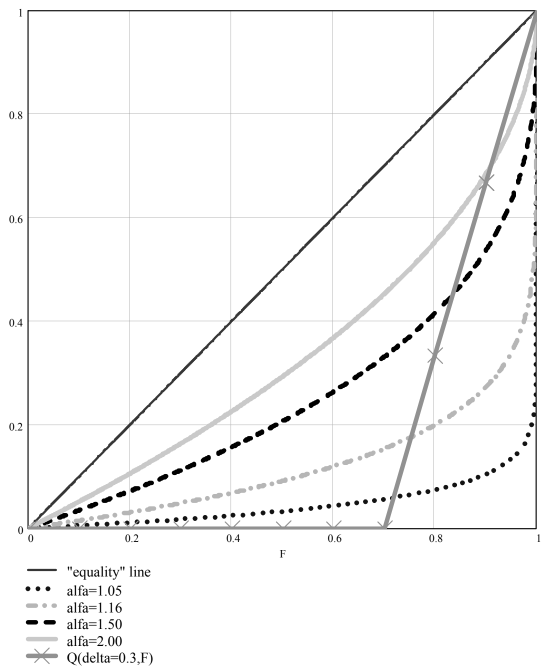

Figure 1. The set of Lorenz curves

Since a mobile terminal of a subscriber may be used by some other people, for example, subscriber's family members, colleagues, friends, and so on, the distribution of 2G services between individuals is more even than the distribution of 2G terminals. Note that in most cases this distribution corresponds to the Pareto law and in order to solve the problem of determination of a number of prosperous users it is necessary to find a transition from the distribution curve of subscribers $Q(\Delta, F)$ to the distribution curve of users $Q(\alpha,F)$. In other words, in order to find the Pareto parameter α and for the expression (2.5) it is reasonable to approximate the distribution (2.7) depending on the parameter Δ by the function (2.6) corresponding to the Pareto distribution with the parameter α where the parameter α satisfies the following requirement

$$\alpha_{opt} = \arg\min d(Q(\Delta, F); Q(\alpha, F)) . \tag{2.8}$$

Here, $d(Q(\Delta, F); Q(\alpha, F))$ is the distance between the approximated function and the approximating one.

As a measure of the distance between these two functions the difference module between the Gini coefficients relating to the functions $Q(\Delta,F)$ and $Q(\alpha,F)$ respectively is considered in [P1]. The desired solution has the following view

$$\alpha_{opt} = \frac{0.5(2-\Delta)}{1-\Delta} . \tag{2.9}$$

When $\alpha = \alpha_{opt}$ the function $Q(\alpha,F)$ has the same value of the Gini coefficient as the function $Q(\Delta,F)$.

Thus, using the expressions (2.5) and (2.9) it is possible to find the relative number of the prosperous users of 2G services. In the subgroup of prosperous users the penetration level of 2G services T exceeds the average value of the 2G penetration level relating to all 2G users Δ in k times. Let us call the coefficient k as the excess coefficient henceforth. Now it is necessary to determine the interval of values of the excess coefficient k while the prosperous 2G users become the potential users of 3G services.

2.3 The choice of the excess coefficient values

It was argued in [27] that the process of the penetration of telecommunication services in time for a group of individuals corresponds to the logistic law. The logistical function is given as

$$y = \frac{B}{1 + e^{[-(t-t_0)/V]}}, \tag{2.10}$$

where B is the saturation level, t_0 is the inflection point of logistical function, $V = 1/B$.

The view of the logistical function with parameters $B = 1$, $V = 1$, $t_0 = 0$ is shown in Figure 2.

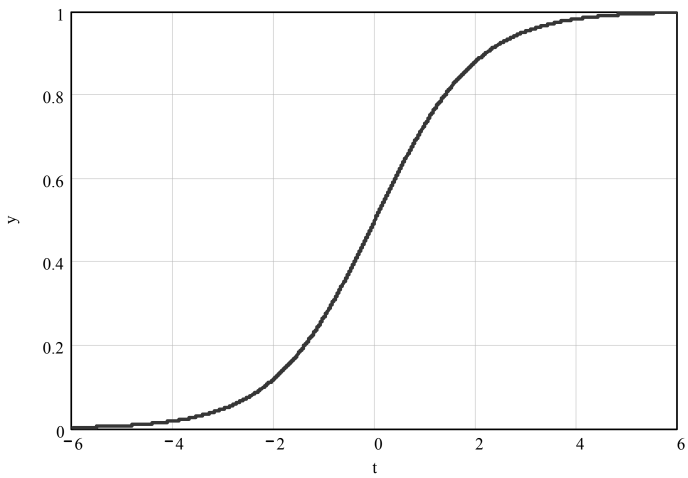

Figure 2. The logistical function

In accordance with the logistic law the process of service penetration in time may be divided into three parts, namely, the first part is exponential growth, the second part is the linear growth, and the third part is saturation.

From the viewpoint of economic principles network operators should deploy the equipment that is able to support new services at the time moment while the greatest demand for the current service. It is so called time for a new attack in marketing. It takes place while the current service development is getting to the second part of linear phase (see Figure 2). In other words, the members of the subgroup of the prosperous users become the potential users of 3G services at the moment when the penetration level of 2G services in the subgroup T achieves 60-80%, i.e. when $k = (0.6 \ldots 0.8)/\Delta$.

This conclusion is confirmed by the relationship between the relative number of potential 3G users and the penetration level of 2G services presented in Figure 3. The relationship is arranged in accordance with (2.5) and (2.9) when $T = 0.8$. As it is seen from Figure 3 while the penetration level of 2G services is 0.6…0.8 the dramatic growth of the relative number of potential 3G users is observed. Note that while the 2G penetration level relating to all 2G users Δ is coming to the 2G penetration level in the subgroup of the prosperous users T then values of the excess coefficient k is tending to 1. In this case, there is a gradual growth of the relative number of potential 3G users depending on values of the Pareto parameter α only. For instance, if the 2G penetration level Δ is more or equal to 0.8 (see Figure 3) then the excess coefficient k is always equal to 1 (coefficient k can not be less than 1) and the further growth of the relative number of potential 3G users becomes more slow. So, when $k = 1$ the expression (2.5) transforms into $R(\alpha)=((\alpha-1)/\alpha)^{\alpha}$. It is seen from the formula that with increasing α the relative number of potential 3G users $R(\alpha)$ is raised, in particular, when $\alpha \to \infty$ then $R(\alpha) \to 1$ that corresponds to the absolutely equal distribution of services between population.

Finally, the absolute value of the number of 3G potential users is obtained as

$$N_{3G} = R(k,\alpha)N\,, \tag{2.11}$$

where N is the number of population in a region.

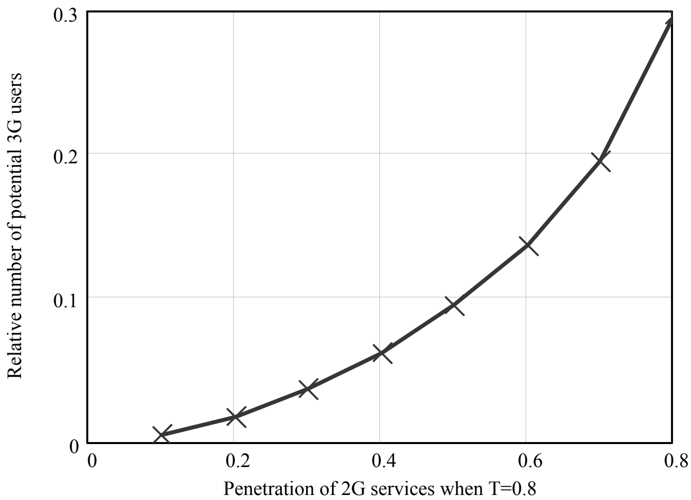

Figure 3. The relationship between the relative number of potential 3G users and the penetration level of 2G services

So, it is quite easy to approximately estimate with help of the expressions (2.5), (2.9), and (2.11) the initial number of prospective 3G users before deployment of 3G wireless systems in a region if the penetration level of 2G services in the region is known.

2.4 Case studies

In this section two simple examples illustrating the proposed approach are considered. The first example is follows. According to statistical information (source: Russian Federal State Statistics Service) there was 8.5 million of population ($N = 8.5 \cdot 10^6$) in Moscow (Russia) and its suburbs in 2002. The penetration level of 2G services in the region was not very high $\Delta = 0.55$. Using the expressions (2.5), (2.9), (2.11) it is possible to estimate both the relative number of potential 3G users in the city as $R(1.1;\ 1.6) = 0.18$ and the absolute number of potential 3G users as $N_{3G} = 1530000$.

Let us consider a region with high level penetration of 2G services, for example, Finland where the 2G penetration level in 2002 was $\Delta = 0.80$ (source: Statistics Finland). The population of Finland in 2002 was 5.2 million. In this case the excess coefficient k is coming to 1 and the relative number of prospective users depends on the Pareto parameter only (see explanation above). Thus, $R(1;3) = 0.3$ and the absolute number of prospective 3G users is N_{3G} = 1560000.

As it is seen from these two examples in these regions there are enough potential users in order to deploy 3G equipment and develop 3G services.

It is worthwhile to make the concluding remark in the end of the chapter that some forecasts of a demand for telecommunication services on a basis of the Pareto law were made in [26,27] in 1994 and 1997. In particular, the telephone density and the number of Integrated Services Digital Network (ISDN) users in some countries were predicted. Today the results of the forecasts seem quite plausible. The prediction error is not more than a few percent. So, for example, the very low demand for ISDN was predicted, in particular it was estimated that the number of ISDN users even in developed countries cannot exceed 5-6% from the number of Public Switched Telephone Network (PSTN) users that corresponds with the real situation.

2.5 Summary

In this chapter the prediction problem of a demand for 3G services in a region has been considered. In particular, the solution method of the problem using the Pareto law, the Lorenz curves, the Gini coefficient and some economic characteristics was proposed. Besides, a novel approach for estimating the Pareto parameter was considered. If the number of population in a region and the penetration level of 2G services in the region are known then the method enables estimating the initial number of prospective 3G users before 3G equipment deployment.

3 ESTIMATING PARAMETERS OF 3G DATA TRAFFIC

The second problem of the 3G network planning is the prediction of traffic parameters. The solution of the problem allows estimating a load on the 3G network equipment. The prediction of data traffic characteristics becomes more complicated when planning the 3G wireless systems because of the following factors.

The 3G wireless systems are characterized by a large diversity of packet switched services that lead to essential quantitative and qualitative changes of data traffic. In particular, a dramatic rise is expected in the rate of transactions and in the amount of data transferred during the transactions. Furthermore, it is important to take into consideration self-similarity that may occur in the 3G wireless networks due to the high inconstancy of burstiness of multiservice traffic [5-8] (see explanation in the beginning of the next chapter).

In this chapter the estimation method for the main parameters of the data traffic, generated by users of 3G services on a level of the PS CN domain, taking into account the above-mentioned 3G features is considered. The method is based on the probabilistic model of events initiated by 3G calls that is formed in accordance with both decomposition of a set of 3G services into some subsets and distribution of potential 3G users into some subgroups. *The method is described in detail in* [P2]. Both the resulting table of the initial data and the resulting table of the expressions for estimating 3G data traffic parameters are presented in the end of the chapter before the case study.

3.1 The decomposition of a set of 3G services into some subsets

The first step of forming the probabilistic model is to make decomposition of a set of 3G services into some subsets. It is supposed that services belonging to one subset generate approximately equal traffic intensity. As criteria for the decomposition an amount of transactions in busy hour and an amount of data transferred during a transaction are used. The decomposition of a set of 3G services is fulfilled into three subsets.

The first subset of services (i = 1) deals with transfer of a short amount of information (E-mails, Web-pages, chats, and so on). These services are characterized by integrity of information blocks. The subset including such services is named “pages”. An average volume of transferred data per transaction for this type of services is quite low, from tens to hundreds of kB.

The second subset (i = 2) includes a transfer of text, color pictures, graphics and so on, for instance, transfer of jpeg, gif, doc, pdf, zip files. The subset is named “pictures”. A demand for the services and an average amount of transferred information per transaction are characterized as the mean level, hundreds of kB.

The third subset (i = 3) comprises multimedia services, which deal with a huge amount of information, for instance, a transfer of mp3, mpeg, avi files. The subset is called “multimedia” and an amount of transferred information for the subset may achieve several tens and hundreds of MB.

The initial information concerning a specific (per each subset of services) distribution of the total amount of transactions in busy hour (γ_i, i = 1,2,3) is used as the numerical criterion for the decomposition of a set of 3G services into some subsets, $\gamma_1 + \gamma_2 + \gamma_3 = 1$.

Note that 3G services within each subset generate the current traffic intensity that has a variance of parameter values of substantially smaller than one relating to the traffic intensity of the total set of 3G services.

3.2 The distribution of 3G users into some subgroups

The second step of forming the probabilistic model is to fulfill the distribution of 3G users into some subgroups since the inequality of a demand for 3G services from different subsets will take place between 3G users. This is because of the fact that a demand for 3G services depends on both a solvency of users and service tariffs, and tariffs on services from the different subsets will be unequal.

All 3G users are divided into three subgroups in accordance with their demand for 3G services from the different subsets defined earlier. The inequality of a demand for each subset of services is approximated by the Pareto law with the parameters (α_i, i = 1,2,3). As a measure of the inequality of the service distribution for each subset the Gini coefficients ($K_{G,i}$, i = 1,2,3) are given and considered as the input data [20,21]. Note that $\alpha_i = 0.5(K_{G,i} + 1)/K_{G,i}$, i = 1,2,3 [24].

Values of the Gini coefficients may be defined on a basis of statistical information and marketing research regarding a demand for 3G services. If there is no information about values of the Gini coefficient then it is possible to extend and develop the approach for estimating the Pareto parameters proposed in the previous chapter.

It is obvious that the least inequality of the distribution takes place for the “pages” subset and the largest one takes place for the “multimedia” subset.

As a rule for the distribution of users into three subgroups (J = 3) the following one described below may be applied. Users from the third (j = 3) subgroup produce 90% of a demand for 3G services from the “multimedia” subset. The subgroup is named “rich”. Users from both the third subgroup and the second subgroup create 90% of a demand for 3G services from the “pictures” subset. The second (j = 2) subgroup is named “middle”. All the rest of users belong to the first (j = 1) subgroup named “poor”.

Using the Lorenz curves (2.6) it is possible to get the relative number of users in each of the subgroups (F_3, F_2, F_1) in accordance with their demand for 3G services from the subsets of 3G services defined in the previous section.

It should be emphasized that that there are not special restrictions for both a choice of the number of subsets of 3G services (I) and a choice of a number of subgroups of users (J). So, for example, in [P3] the present method is developed to estimate the main parameters for symmetrical and asymmetrical traffic separately. In the method, other principles of decomposition of services are proposed. In the first step of the decomposition, a set of services is divided into two subsets. The first subset includes services concerning the real-time establishment of connectivity between endpoints. It is characterized by the transfer of the symmetrical traffic and the strict control of QoS. The second subset comprises of such services that generate the asymmetrical traffic. After that, each of the above-mentioned subsets of services is divided into three classes in accordance with features of the generated traffic intensity. The division of the first subset of services creating symmetrical traffic into three classes is fulfilled taking into account the flow rate initiated by users. The division of the second subset of services initiating the asymmetrical traffic into three classes is fulfilled taking into account an amount of information (on average) initiated by users. Such kind of decomposition enables estimating separately the symmetrical and the symmetrical load on network equipment.

3.3 Forming the probabilistic model of events initiated by 3G calls

In accordance with the rules of the decomposition of 3G services and the distribution of users of these services the probabilistic model of the initiation of transactions based on an intersection of events from two statistically independent exhaustive classes is formed.

The events included in the first class correspond with demands for services from the "pages", "pictures", "multimedia" subsets respectively and denoted by index i= 1,2,3. The events included in the second class correspond with demands initiated by users from the "poor", "middle" and "rich" subgroups respectively and denoted by index j = 1,2,3. It is possible to arrange nine intersections of events from these two classes (i,j: i = 1,2,3; j = 1,2,3). The first event (i = 1, j = 1) is that user from the first subgroup (j=1) initiates a demand for a service from the first subset (i = 1). At the second event (i = 2, j = 1) user from the first subgroup initiates a demand for a service from the second subset. At the third event (i = 3, j = 1) user from the first subgroup makes a demand for a service from the third subset. The fourth event (i = 1, j = 2) is that user from the second subgroup makes a request for a service from the first subset and so on.

Thus, the probabilistic model of the events allows segregating nine segments from the common flow of transactions initiated in busy hour by support procedures of such services. The variance of parameter values of each random flow is less than the one relating to the common flow.

3.4 The approach for estimating the specific rate of transactions and the traffic intensity

After the probabilistic model has been composed the problem is to determine the specific (per user) rate of transactions (λ_{ij}) in busy hour for nine (i,j = 1,2,3) intersections of events from the two above-mentioned classes. The approach for estimating values of the parameter is presented in detail in [P2]. It is based on solving system of three equations that are formed for each subset of 3G services. Each system of the equations is assigned on a basis of the share of transactions (β_{ij}) in busy hour relating to users of the j-th subgroup when services from the i-th subset are initiated, i,j=1,2,3. The parameters β_{ij} may be expressed by two different ways that give a possibility to form the required equations.

The first way is to define the parameter β_{ij} on a basis of the inequality of a demand for 3G services. Values of the share of transactions β_{ij} are calculated for each of nine events using values α_i, i = 1,2,3 and F_j, j = 1,2,3 obtained earlier. So, the parameter values β_{ij} are known. The second way is to express the parameter β_{ij} over the specific (per user) rate of transactions in busy hour λ_{ij} (i,j = 1,2,3). It gives a possibility to formulate the above-mentioned system of the equations for each subset of 3G services with unknown parameters λ_{ij} (i,j = 1,2,3). The system may be presented in the matrix-vector form as follows

$$\begin{pmatrix} (\beta_{i1}-1)F_1 & \beta_{i1}F_2 & \beta_{i1}F_3 \\ \beta_{i2}F_1 & (\beta_{i2}-1)F_2 & \beta_{i2}F_3 \\ \beta_{i3}F_1 & \beta_{i3}F_2 & (\beta_{i3}-1)F_3 \end{pmatrix} \begin{pmatrix} \lambda_{i1} \\ \lambda_{i2} \\ \lambda_{i3} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \quad (3.1)$$

The system (3.1) has three one-parameter families of solutions that are given in [P2]. In accordance with solutions the rate of transactions in busy hour for 8 segments of data traffic is determined λ_{ij} ($i=2,3;$ j = 1,2,3). Usually, the parameter value λ_{11} is known on a basis of statistical information concerning a demand for 2G services. For this reason it is worthwhile to add the value λ_{11} to the input data to concretize these solutions.

After that it is possible to get the traffic intensity (the load) on nodes of the PS CN domain. The traffic intensity depends on both the obtained values of rate of transactions λ_{ij} (i,j = 1,2,3) and the average duration of a transaction T_i (i = 1,2,3). In order to determine the parameter T_i (i = 1,2,3) it is worthwhile to add to the input data the value of the average amount of information transferred during a transaction w_i (i = 1,2,3) and the value of the average rate of transferred data in a radio-channel c on the interface Mobile Station - Base Station (Node B).

Then, the following data traffic parameters may be estimated

- The specific (per user) traffic intensity generated by calls from the i-th subset of 3G services that are initiated by users from the j-th subgroup (S_{ij});
- the traffic intensity created by calls from the i-th subset of 3G services that are initiated by all 3G users (S_i);
- the total traffic intensity on nodes of the PS CN domain (S).

Besides, the amount of information transferred during a month (per user) when services from the i-th subset are requested by users from the j-th subgroup (U_{ij}) and the amount of information transferred during a month when services from the i-th subset are requested by all 3G users (U_i) may be calculated.

Below both the Table II of the initial data and the Table III of the expressions for estimating 3G traffic parameters are presented. Sequential realization of the expressions gives the method for estimating the main parameters of data traffic generated by support procedures of 3G services.

TABLE II. INITIAL DATA FOR THE ESTIMATING 3G DATA TRAFFIC PARAMETERS

PARAMETER / INDEX	MEANING
$i = 1,2,3,\dots I$	the index of a subset of 3G services
$j = 1,2,3,\dots J$	the index of a subgroup of 3G users
γ_i	the specific (per each subset of services) distribution of the total amount of transactions in busy hour
$K_{G,i}$	the Gini coefficient
λ_{11}	the specific (per user) rate of transactions in busy hour for the event when users from the “poor” subgroup initiate a demand for services from the “pages” subset
w_i, kbit/trans	the average amount of information transferred during a transaction
c, kbit/s	the average rate of transferred data in a radio-channel
N_{3G}	the number of 3G users
k	the concentration factor showing a portion of daily transactions in busy hour

TABLE III. EXPRESSIONS FOR ESTIMATING 3G DATA TRAFFIC PARAMETERS

EXPRESSIONS	MEANING
$\alpha_i = \dfrac{0.5(K_{G,i}+1)}{K_{G,i}}$	the parameter of the Pareto distribution
$F_3 = 0.9^{\frac{\alpha_3}{(\alpha_3-1)}}$	the relative number of users in the "rich" subgroup
$F_2 = 0.9^{\frac{\alpha_2}{(\alpha_2-1)}} - F_3$	the relative number of users in the "middle" subgroup
$F_1 = 1 - F_2 - F_3$	the relative number of users in the "poor" subgroup
$\beta_{i1} = 1-(1-F_1)^{\frac{(\alpha_i-1)}{\alpha_i}}$	the share of transactions in busy hour relating to users of the "poor" subgroup when services from the *i*-th subset are initiated
$\beta_{i2} = 1-(1-F_1-F_2)^{\frac{(\alpha_i-1)}{\alpha_i}} - \beta_{i1}$	the share of transactions in busy hour relating to users of the "middle" subgroup when services from the *i*-th subset are initiated
$\beta_{i3} = 1-\beta_{i1}-\beta_{i2}$	the share of transactions in busy hour relating to users of the "rich" subgroup when services from the *i*-th subset are initiated
$\lambda_{12} = \dfrac{\lambda_{11}\beta_{12}F_1}{\beta_{11}F_2}$	the specific (per user) rate of transactions in busy hour for the event when users from the "middle" subgroup initiate a demand for services from the "pages" subset
$\lambda_{13} = \dfrac{\lambda_{11}\beta_{13}F_1}{\beta_{11}F_3}.$	the specific (per user) rate of transactions in busy hour for the event when users from the "rich" subgroup initiate a demand for services from the "pages" subset
$\lambda_{ij} = \dfrac{\gamma_i\beta_{ij}\sum_j \lambda_{1j}F_j}{\gamma_1 F_j}; i=2,3; j=1,2,3$	the rest two family of solutions of the system (3.1) for "pictures" and "multimedia" subsets
$L_i = \sum_j F_j\lambda_{ij}$	the specific (per user) rate of transactions in busy hour for each subset of services
$T_i = \dfrac{w_i}{c}, s$	the average duration of a transaction
$U_{ij} = \dfrac{30\lambda_{ij}w_i}{8k}$	the amount of information transferred during a month (per user) when services from the *i*-th subset are requested by users from the *j*-th subgroup
$U_i = N_{3G}\sum_j F_j U_{ij}$	the amount of information transferred during a month when services from the *i*-th subset are requested by all 3G users
$S_{ij} = \dfrac{\lambda_{ij}T_i}{3600}, Erl$	the specific (per user) traffic intensity generated by calls from the *i*-th subset of 3G services that are initiated by users from the *j*-th subgroup
$S_i = \dfrac{N_{3G}T_i\sum_j F_j\lambda_{ij}}{3600}, Erl$	the traffic intensity created by calls from the *i*-th subset of 3G services that are initiated by all 3G users
$S = \dfrac{N_{3G}\sum_i T_i\sum_j F_j\lambda_{ij}}{3600}, Erl$	the total traffic intensity

The presented expressions have been obtained when both the number of subsets of 3G services I and the number of subgroups of 3G users J are equal to three. However, as it was emphasized above, all derived expressions may be easily transformed when values I and J are different from three. Note that the number of 3G users (N_{3G}) may be approximately estimated on a basis of the approach proposed in the previous chapter.

3.5 Case study

Let us apply the proposed methods for estimating parameters of data traffic coming to/from Packet Data Serving Node (PDSN). PDSN is the main node of the network subsystem of IMT-MC (International Mobile Telecommunications-2000 Multi-Carrier) [33]. IMT-MC-450, sometimes called CDMA2000 1X or IS-95C, is a 3G system (using the 450 MHz frequency band) that allows for mobile Internet access, high-speed data and image transmission, and high quality voice transmission protected from unauthorized access [34]. On the first step of equipment deployment of the system IMT-MC-450 (phase IMT-MC 1X) the data traffic parameters have been estimated for one of network operators providing mobile services to users in NMT-450 standard.

The initial values are given in the Table IV and the calculation results are presented in the Table V. Note that in the Table IV values of the Pareto distribution are given right away as the initial data.

TABLE IV. INITIAL VALUES FOR THE CASE STUDY

PARA-METER	PARAMETER NAME	"PAGES" SUBSET	"PICTURES" SUBSET	"MULTIMEDIA" SUBSET
α_i	Pareto parameter	3.5	1.5	1.02
γ_i	specific distribution of trans. in busy hour (BH) for each subset of services	0.9	0.085	0.015
w_i, kbit/trans	av. amount of information per transaction	1	100	10000
N_{3G}	number of potential 3G users	45000		
c, kbit/s	av. data rate in a radio-channel	50		
λ_{11}, trans/user	specific rate of transactions in BH for the fist segment	0.5		
k	concentration factor	0.1		

TABLE V. ESTIMATED VALUES OF 3G DATA TRAFFIC PARAMETERS

PARA-METER	PARAMETER NAME	"PAGES" SUBSET	"PICTURES" SUBSET	"MULTIMEDIA" SUBSET
F_3	relative number of "rich" users	0.005		
F_2	relative number of "middle" users	0.724		
F_1	relative number of "poor" users	0.271		
β_{ij}	share of transactions in BH for each defined segment	$\beta_{11} = 0.202$ $\beta_{12} = 0.776$ $\beta_{13} = 0.022$	$\beta_{21} = 0.100$ $\beta_{22} = 0.733$ $\beta_{23} = 0.167$	$\beta_{31} = 0.006$ $\beta_{32} = 0.094$ $\beta_{33} = 0.900$
λ_{ij}	specific (per user) rate of transactions in BH for each defined segment	$\lambda_{11} = 0.5000$ $\lambda_{12} = 0.7186$ $\lambda_{13} = 3.1126$	$\lambda_{21} = 0.0234$ $\lambda_{22} = 0.0641$ $\lambda_{23} = 2.2767$	$\lambda_{31} = 0.0003$ $\lambda_{32} = 0.0014$ $\lambda_{33} = 2.1682$
L_i	specific (per user) rate of trans. in BH for each subset of services	0.67	0.02	0.01
T_i, *s*	average transaction duration	0.02	2	200
U_{ij}, kB	amount of data transferred during a month (per user) for each defined segment	$U_{ij} = 18.75$ $U_{ij} = 26.95$ $U_{ij} = 116.72$	$U_{ij} = 87.62$ $U_{ij} = 240.36$ $U_{ij} = 8537.6$	$U_{ij} = 95.54$ $U_{ij} = 542.75$ $U_{ij} = 813067.5$
U_j, MB	amount of data transferred during a month for each subset of services	1131.40	10685.44	188566.5
S_{ij}, Erl	traffic intensity for each defined segment	$S_{ij} = 0.034$ $S_{ij} = 0.130$ $S_{ij} = 0.004$	$S_{ij} = 0.158$ $S_{ij} = 1.161$ $S_{ij} = 0.264$	$S_{ij} = 0.173$ $S_{ij} = 2.621$ $S_{ij} = 25.142$
S_i, Erl	traffic intensity for each subset of 3G services	0.168	1.583	27.936
S, Erl	total traffic intensity	29.686		

The calculations have been made using the expressions from the Table III.

The calculation results show that on the first step of deployment of the IMT-MC-450 equipment the expected data traffic on PDSN may by characterized by the following most important parameters

- the total traffic intensity in busy hour is about 30 Erl;
- the rate of transactions per user in busy hour is 0.67 trans/user for the subset of services named "pages", 0.06 trans/user for the "pictures" subset, 0.01 trans/user for the "multimedia" subset;

- the amount of user data coming to/from PDSN during a month is more than 1100 MB for the "pages" subset of services, more than 10000 MB for the "pictures" subset, and more than 1800000 MB for the "multimedia" subset;
- the average duration of a transaction is 0.02 s for the "pages" subset of 3G services, 2 s for the "pictures" subset and 200 s for the "multimedia" subset.

It is seen from the results that services from the "pages" subset generate a small part of the traffic intensity on network elements. However, they initiate the largest number of transactions. At the same time, services from the "multimedia" subset give the main part of the traffic intensity, but have the least number of transactions. This fact indicates that both the number of initiated transactions and the amount of transferred or received data should be taken into account when forming tariffs on 3G services.

3.6 Summary

The estimation method for the main parameters of 3G user traffic has been considered. The method is based on decomposition of all services into three subsets and on the approximation of distribution of 3G users separately in each subset. With help of the proposed method the following main parameters of 3G data traffic may be estimated:

- the specific (per user) rate of transactions in busy hour for each of the defined traffic segments;
- the average duration of 3G calls for each subset of 3G services;
- the amount of user data coming to/from packet nodes of the network subsystem during a month;
- the distribution of traffic intensity generated by 3G calls in each subset of 3G services and in each subgroup of 3G users;
- the total traffic intensity on packet nodes of the network subsystem.

The calculation results show that it is reasonable to impose tariffs on 3G services depending on the number of initiated transactions and on the volume of transferred/received information.

4 PERFORMANCE EVALUATION OF IM-SUBSYSTEM ELEMENTS

The 3G system is in evolution through new phases. The common trend has many different names, for example, “Mobile IP”, “3G all IP”, “End-to-End IP”. From a 3G point of view, a full-scale IP implementation is defined as one targeted phase of the 3G development path [1]. There are different approaches towards the global cellular system, 3G. These approaches are summarized in the Table VI taken from [1].

TABLE VI. 3G VARIANTS AND THEIR BUILDING BLOCKS

VARIANT	RADIO	SWITCHING	2G BASIS
3G (US)	WCDMA, EDGE, CDMA 2000	IS-41	IS-95, GSM 1900, TDMA
3G (Europe)	WCDMA, GSM, EDGE	Advanced GSM NSS and packet core	GSM900/1800
3G (Japan)	WCDMA	Advanced GSM NSS and packet core	PDC

Within this thesis Universal Mobile Telecommunication System (UMTS) as a third generation platform for mobility and services (the second variant in the Table VI) is considered. UMTS as a part of the IMT-2000 standards family is one of the most important European Telecommunications Standards Institute (ETSI) projects in creating the mass market for high-quality wireless multimedia communications [36]. There are some implementation of UMTS defined by the 3G Partnership Project (3GPP) and ETSI and called releases. So, UMTS Rel’ 99 offers the same services as that of GSM Phase2+. In UMTS Rel’ 4 mechanisms and arrangements for multimedia are implemented [37]. In UMTS Rel’ 5 the further evolution to all IP takes place and the transport network utilizes IP networking as much as possible. The evolution process is characterized by the introduction of the additional subsystem named the Internet Protocol Core Network Subsystem (IM-subsystem) in UMTS Rel’ 4,5,6 [38,39]. The main purpose of the subsystem is to offer uniform methods to perform VoIP (Voice over IP) calls and to support IP based multimedia services. The IM-subsystem covers all PS CN domain elements to provide both VoIP and multimedia services [39].

The problem of performance evaluation of IM-subsystem elements is one of the main problems of 3G core network planning. This is because of the fact that the traffic generated by 3G services may be self-similar or long-range dependent in nature (i.e., bursty over a wide range of time scales).

Self-similarity is observed in different networks; in particular, in local area networks [40], Internet [41], wireless networks [5] and others. It is shown in [7] that in GPRS in the case of aggregated traffic and also in the case of individual WAP and WEB traffic traces, the results strongly suggest long-range dependency (values of the Hurst parameter are about 0.8). Besides, the packet arrival process of WAP and WEB traffic may be considered as a class of processes consisting of the superposition of an infinite number of ON/OFF-sources. Through the characterization of the sum of the covariances, it is possible to establish a simple explicit necessary and sufficient condition for the process to be long-range dependent [42]. It is reasonable to suppose that self-similarity may occur in 3G wireless networks as well. This is in sharp contrast to commonly made traffic modeling assumptions, because self-similarity is characterized by stronger dependence of a variance from time than linear dependence [7]. The traffic does not smooth out in the case of aggregation, leading to congestion situations and packet-drops due to the burstiness of the traffic.

In the case of self-similar traffic the applied methods for performance analysis and network dimensioning are different from those applied to statistically more simple traffic, which can be modeled with Markovian processes [43,44]. For example, the queue tail behavior is heavy-tailed in the case of self-similar input traffic [25].

Thus, the use of the classic teletraffic theory for performance evaluation of packet multiservice network elements gives essential faults, in particular, network parameters may be underestimated [41,45,46]. In literature [23-25] the approaches of overcoming such sort of difficulties are considered. In our research the results from [24] concerning self-similar multiservice traffic are developed and applied for performance evaluation of such UMTS Rel'5 IM-subsystem element as Gateway GPRS Support Node (GGSN). The method for GGSN performance evaluation is based on using the FBM/D/1/W queueing system. *It described in detail in* [P4].

4.1 UMTS Rel' 5 core network architecture aspects

The reference architecture for UMTS Rel' 4 and Rel' 5 from 3GPP TR 23.821 is the same [1]. In the development of UMTS Rel' 5 the focus has shifted to the PS CN domain, which has been extended with IM-subsystem functionality. It was mentioned in Press Release (06.04.2005) of Global mobile Suppliers Association (GSA) that IM-subsystem functionality in UMTS Rel' 6 is developed to support inter-working with circuit-switched networks, non-IM-sybsystem networks and 3GPP2 based CDMA systems. The detailed architecture of UMTS Rel'5 core network based on configurations [38] with denoted IM-subsystem elements is presented in [P4].

Let us consider the vision of UMTS Rel'5 from the All IP point of view taken from [1] is shown in Figure 4. As seen from Figure 4, the principle of allocation of data flows between end users and GGSN leads to increasing of the load on the network elements while approaching to GGSN. It is obvious that GGSN is a node that may be under the influence of self-similarity in UMTS. The most important events determining the load on GGSN on the network level are arriving IP packets. Currently, a transport technology for delivery of IP packets to/from GGSN is not defined uniquely. For instance, ATM may be applied as one of the possible cases of such technology [46].

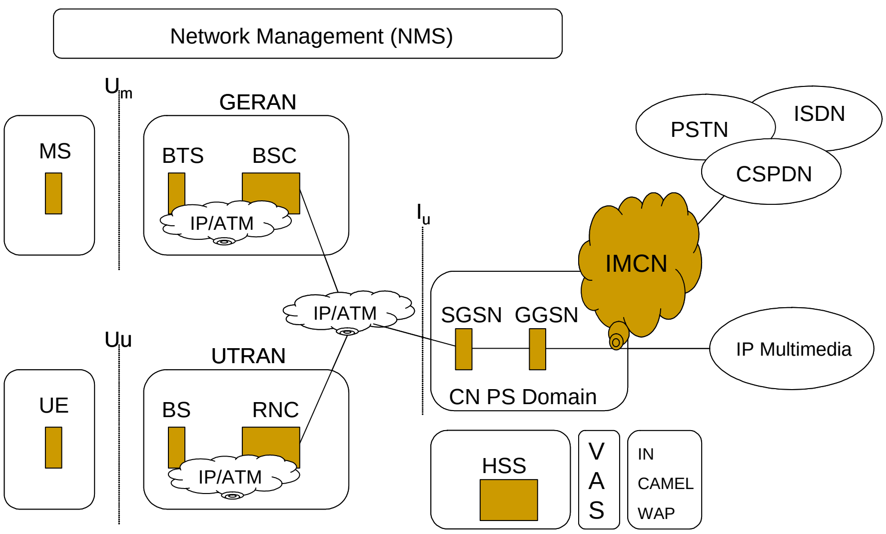

Figure 4. Vision of UMTS Rel' 5 (All IP)

4.2 The traffic model and the queueing system model

It is assumed that values $s(t)$ of a random process with interdependent increments are the total load arriving to the node (server) in the time point $t > 0$. Current values $s(t)$ may be determined by the number of the information units (bytes, ATM cells, IP packets, and so on, depending on a level of protocol stacks). If the corresponding process is ordinary then an increment is one information unit. The intensity of increments is the rate parameter λ, 1/s. Realizations of process $s(t)$ are non-decreasing step functions with increments taking place in random time points.

A sample of a random process $s(t)$ is a random variable $S_T = s(t = T), T \geq 0$ that may be approximated by a Gaussian random variable [47]

$$S_T = \lambda T + \sqrt{b(T)} \cdot x, \tag{4.1}$$

where $x = N(0,1)$ is the normalized Gaussian random variable with the zero mean and the unit variance, $b(T)$ is the variance of S_T.

If $b(t) = \sigma^2 t$ and $t > 0$ then the one-dimensional distribution of probabilities of the process $s(t)$ coincides with the corresponding distribution of the Brownian motion process. There are different ways to model the self-similarity load [48,49]. With reference to (4.1) self-similarity is taken into account as

$$b(T) = (\sigma^2 T)^{2H}, \qquad 0.5 \leq H < 1, \tag{4.2}$$

where H is the Hurst parameter. The Hurst parameter is a measure of the degree of self-similarity of the aggregate traffic stream. Expressions (4.1) and (4.2) give the model of the common traffic intensity arriving to the server input by a time point $t = T$.

It assumed that this server is modeled by queueing system with deterministic service rate C, 1/s and the buffer size (W-1), $1 \leq W < \infty$. The queueing system is the stable one because there is a stationary probability distribution if $C > \lambda$. In accordance with the Kendall's notation for queues the system is G/D/1/W [50]. The corresponding system may be also defined as FBM/D/1/W if the expressions (4.1), (4.2) are fulfilled [24]. Here, the FBM is a normalized fractional Brownian motion, i.e. the corresponding process has self-similar nature.

4.3 The evaluation problem and its solution

Note that the arriving load is a random process. Therefore, an event may occur when the server buffer is overflowed. The probability of the event is defined by statistical properties of unserved traffic process that may be given as

$$V(t) = \max[0, S(t) - Ct]. \tag{4.3}$$

The evaluation problem is to determine values of parameters C and W such that the probability of occurence of unserved load more than the parameter W must not exceed the preset threshold ε:

$$P[V(t) > W] = \varepsilon, \quad t > \varepsilon \quad 0 < \varepsilon << 1. \tag{4.4}$$

Taking into consideration the approximation of the random process $s(t)$ by the random Gaussian variable defined by the expressions (4.1) and (4.2) the bound for the buffer saturation probability has the following view

$$P[V(t) > W] \geq \max_{T>0} P[x > \alpha(T)], \tag{4.5}$$

where $\alpha(T) = [(C - \lambda)T + W]/(\sigma^2 T)^H$.

Since the random variable X is the normalized unbiased Gaussian random one, the expression (4.5) may be transformed as

$$P[V(t) > W] \approx \max_{T>0}[0.5\exp(-\alpha^2(T)/2)]. \tag{4.6}$$

After that, it is possible to get the expression binding the parameters *C, W, λ* and the buffer saturation probability ε

$$-\ln\varepsilon \approx \min_{T>0} \frac{((C-\lambda)T + W)^2}{(\sigma^2 T)^{2H}}. \tag{4.7}$$

Using the approach presented in [24] the solution of the functional equation (4.7) may be found in the following view

$$C/\lambda = 1 + n\left(\frac{\sqrt{-2\ln\varepsilon}W^{H-1}H^H}{(1-H)^{H-1}}\right)^{1/H}, \qquad 0.5 \leq H < 1, \tag{4.8}$$

where $n = \sigma^2/\lambda$.

Substituting *n, W,* λ and ε values in (4.8) it is possible to get the upper bound (if $H = 0.9$) and lower bound (if $H = 0.5$) of the server service rate *C*. Besides, the utilization factor ($\rho = \lambda/C$), sometimes called the server utilization, may be determined as the parameter inversed to the ratio C/λ. This is one of the main parameters characterizing queueing systems. Figure 5 illustrates relationships between the utilization factor and the parameter *n* for various values of the Hurst parameter and the buffer capacity *W* when $\varepsilon = 10^{-7}$.

Trends of the curves (Figure 5) show that it is quite important to take into account the self-similarity influence while assigning server parameters. In Figure 5 the section of the curves when $n = 1$ and $H = 0.5$ corresponds to the case of the Poisson arrival process. According to investigations made in [P5] values of the Hurst parameter for fragments of Internet traffic may achieve values exceeding 0.8 that corresponds to the results obtained in [51]. For these reason the curves when $H = 0.9$ are presented in Figure 5.

The expression for estimating the upper bound of the average queue length in the server buffer is obtained on a basis of results [25,53]. The lower bound of the average queue length (q_{min}) may be estimated on a basis of the classical result for M/D/1 queueing system [50].

$$q_{\max} = \frac{(\lambda/C)^{1/2(1-H)}}{(1-\lambda/C)^{H/(1-H)}}, \qquad q_{\min} = \frac{\lambda/C}{1-\lambda/C} - \frac{(\lambda/C)^2}{2(1-\lambda/C)}. \tag{4.9}$$

The upper and the lower bounds for average service time (τ) are determined using the Little result [47,50]

$$\tau_{\max} = q_{\max}/C, \qquad \tau_{\min} = q_{\min}/C. \tag{4.10}$$

Thus, using the expressions (4.8), (4.9) and (4.10) it is possible to determine bounds for the main performance measures of the single server under the self-similarity load influence.

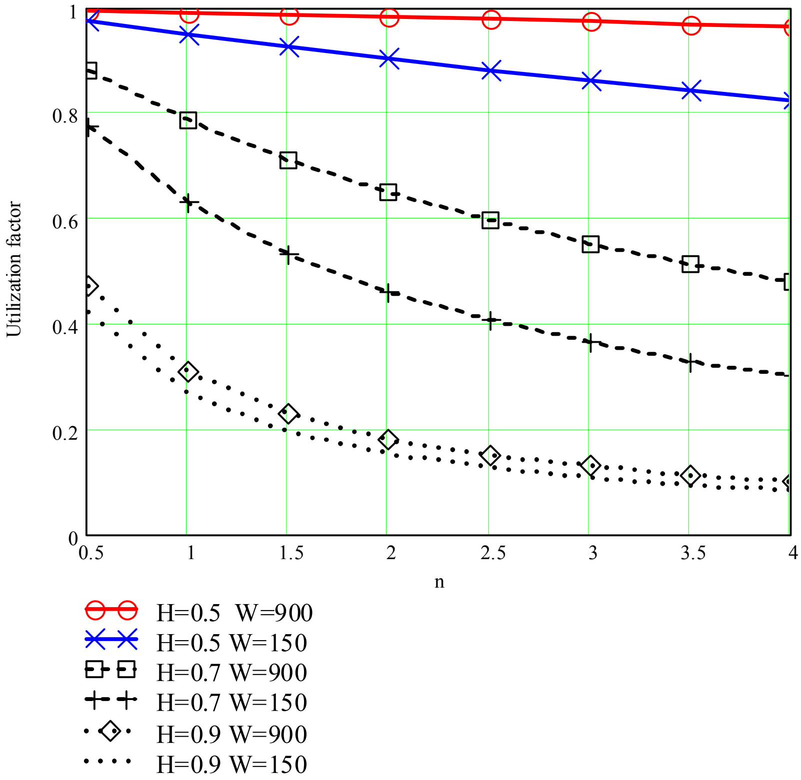

Figure 5. Self-similarity influence on the server utilization factor

4.4 Case study

In this section the example illustrating the above-presented method is considered. It is worth mentioning that there are no strong recommendations on transport network protocols on SGSN-GGSN interface at the present moment. The transport technology could be IP or ATM but this is the operator's choice. In this example, ATM is used as underlying technology for delivery of IP packets. The rate of information units (ATM cells) arriving on SGSN is multiple (k)·of 2 Mbit/sec. The parameters characterizing the server normal functionality may be estimated by the following way.

Let k = 20 and in average 30% of the channel throughput is in use during the messages delivery to SGSN. Then, the value of the intensity of ATM cells arriving on the SGSN input is $\lambda \approx 30000\ s^{-1}$. If a number of SGSNs connected to the GGSN is 4 then the total value of the intensity of ATM cells arriving to GGSN input is $\lambda \approx 120000\ s^{-1}$. In accordance with (4.8), (4.9) and (4.10) the relationships between the GGSN server capacity, the upper bound for average queue length in the GGSN buffer, the upper bound for the average service time of information units in the GGSN buffer and the parameter n $(n = \sigma^2/\lambda)$ are shown in Figures 6, 7, 8 respectively ($W = 50, 200$; $H = 0.8$; $\varepsilon = 10^{-7}$).

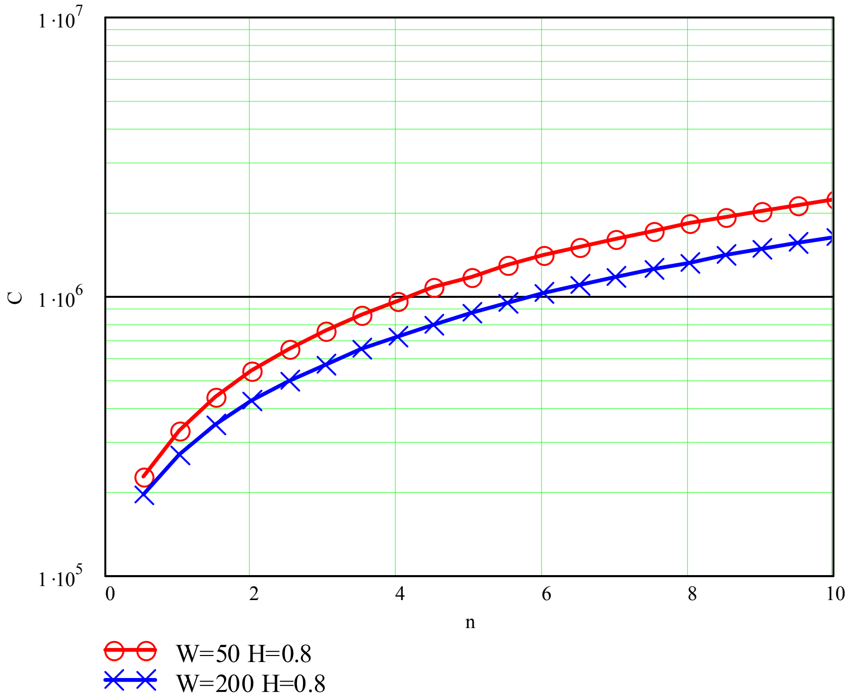

Figure 6. GGSN server capacity estimating

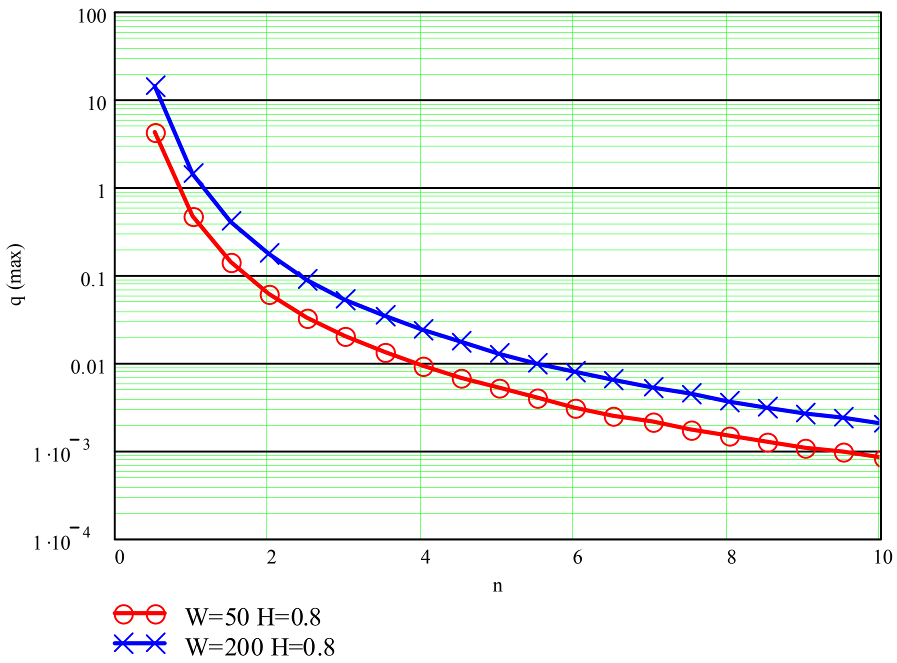

Figure 7. The upper bound for average queue length in the GGSN buffer

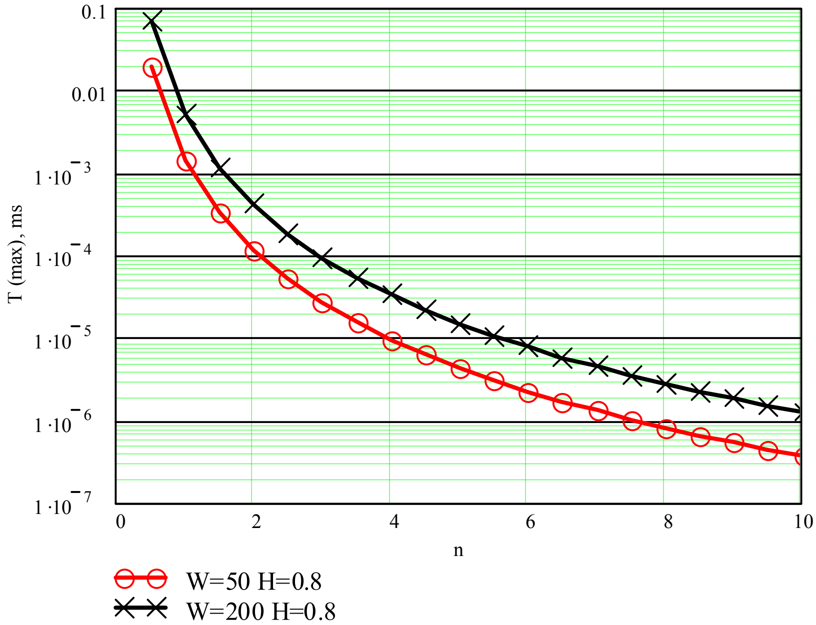

Figure 8. The upper bound for the average service time of information units in the GGSN buffer

4.5 Summary

In this chapter the influence of self-similar input on GGSN performance in UMTS Rel'5 IM-subsystem has been analyzed. FBM/D/1/W queueing system for evaluating the GGSN parameters was applied. The submitted method enables determining the following probabilistic and time characteristics:

- upper and lower bounds for the GGSN service rate;
- upper and lower bounds for the average queue length in the GGSN buffer;
- upper and lower bounds for the average service time of information units in the GGSN buffer;
- the server utilization.

The obtained results point to a need to take into account self-similarity while assigning the GGSN parameters.

As well known, when providing multimedia services based on IP technologies one of the main aspects is to ensure Quality of Service (QoS). From this point of view the presented method may be extended for performance evaluation of other IM-subsystem elements, in particular, for SGSN performance evaluation.

This chapter concludes the first part of the thesis devoted to CN planning aspects for the 3G wireless networks.

5 Reliability and Cost Parameters Estimation for the 4G RAN Ring Topology

The second part of the thesis is devoted to RAN planning aspects for the 4G wireless networks. It consists of the fifth and the sixth chapters. In the fifth chapter the estimation problem of reliability and cost parameters for a new ring RAN topology is considered. The solution method of the problem is given in details in [P6]. The sixth chapter deals with the optimization problem of physical link ring configuration between 4G RAN elements. The method for minimum cost ring configuration of physical links between BSs in 4G RAN is presented in detail in [P7]. Moreover, one more aspect concerning 4G RAN planning is considered in [P8]. This publication is devoted to the universal transport access network planning. Resources of the transport access network may be shared between various switched networks including 4G networks. The specificity of wireless access is taken into account when planning such multiservice access network. The location of 4G RAN nodes such as BSs and Radio Network Controllers (RNCs) may be adjust in accordance with coordinates of points obtained with help of the method presented in [P8].

Thus, we start the second part of the thesis with the estimation of reliability and cost parameters for the 4G RAN ring topology.

5.1 Ring topology definition for 4G RAN

New generation mobile communication systems beyond IMT-2000 (the 4G wireless systems) should provide broadband mobile multimedia services and offer the capacity of 100 Mb/s as the maximum for the data communication channels [15]. In such systems, because there will be a need to deal with the huge amount of traffic the coverage by a single base station is supposed to be shorter than that of the 3G wireless systems. There will be a trend from macro cells to micro cells [13]. With reduced cell size, the 4G RAN will comprise more BSs and utilize links with a capacity more that 23-fold that of those currently used [14]. As a result, more frequent handover will occur resulting in a heavy load on both the BSs entrance links and the RNC signal processing equipment [13]. This may lead to a serious cost increase in the RAN [13].

The 3G RAN is not optimised to high-speed micro-cellular networks. The model of 3G RAN configuration is shown in the left side of Figure 9. It is seen that BSs in such configuration are connected to their dedicated RNC directly. This is known as the radial (tree) topology. In the current 3G releases BSs do not have routing capability, therefore, traffic between them has to be forwarded through the dedicated RNC [54,61].

In [3,13,14] a new RAN topology for the 4G systems from points of view of a load and routing capabilities has been considered and analytically analyzed. This topology is well-known as the ring one. As shown in the right side of Figure 9, in such topology BSs are linked to each other and, the “main” BS connected to the RNC is assigned. BSs in a cluster may be connected to each other by a kind of local area network [13]. The information transportation in the 4G RAN is based on IP protocols [13].

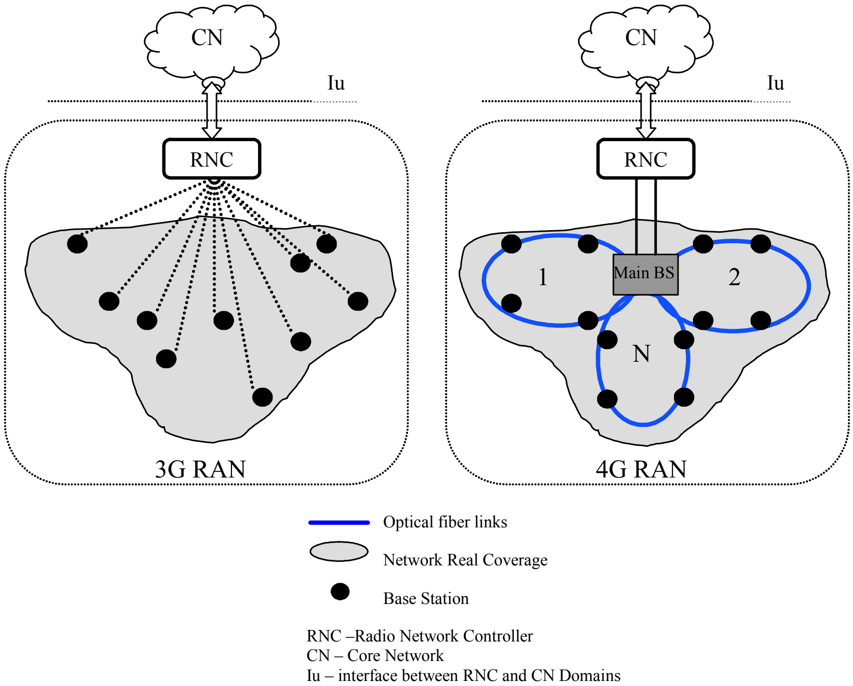

Figure 9. The models of RAN configuration

However, the ring topology should be analyzed not only from the viewpoint of load and routing capabilities. It is very important to consider the topology from the viewpoint of reliability and cost as well. So, the objective of the study is to make the quantitative estimation of reliability and cost parameters of the ring topology, to compare its parameters with those of other topologies and to give recommendations on applying different topologies of physical links to the RAN for mobile communication systems beyond IMT-2000.

5.2 Ring topology reliability estimation

As it is emphasized in [15] optical fiber links should be used as the physical links to connect BSs in the 4G RAN, first of all, from the viewpoint of link capacity. For new generation mobile Radio Access Network new technologies using optical fiber have been proposed in [15, 63] such as fiber and free-space hybrid optical (FFHO) networking and Radio on Fiber (ROF) technology. Note that when employing radio links between the RNC and BSs, a higher frequency band will be used to transmit broadband signals. However, the higher frequency bands suffer from rain or fog attenuation and the signal transmission range is limited to less than a few kilometers [14]. Therefore, radio repeaters are necessary for BSs remote from the RNC that will cause a cost increase in the RAN.

Thus, it is supposed that BSs are connected to each other by optical fibers in the RAN. As the model for 4G RAN configuration the ring topology presented in the right part of Figure 9 is considered.

As the reliability we mean the ability of the 4G RAN to perform its required functions under stated conditions for a given time interval at possible damages which can occur in it.

As a quantitative parameter of the topology reliability, the average number of knocked out BSs is calculated. Such BSs cannot serve its mobile terminals because of damages in the RAN physical links connecting BSs to each other. BS is knocked out if there are damages in the ring on both sides of this BS.

The problem is formulated as follows. The number of rings in the RAN, $N \in \{1,2,\ldots\}$; the number of damages in the RAN, $L \in \{0,1,2\ldots\}$; the number of BSs in a ring, $M \in \{1,2,\ldots\}$ are known. It is supposed that the number of BSs in each ring is equal.

The problem now is to find the average number of knocked out BSs depending on the number of rings in the RAN, the number of BSs in the ring, the damages number in the RAN - $\hat{X}(N,M,L) \in \{1,2,...,z\}$.

The principles of probability theory are used in [P6] to solve the problem.

The common expression for estimating the average number of knocked out BSs is

$$\hat{X}(N,M,L) = \sum_{L=1}^{\infty} \hat{X}(N,M/L) p_L , \tag{5.1}$$

where p_L is the probability of L damages in the RAN, $L \in \{0,1,2\ldots\}$; $\hat{X}(N,M/L)$ is the average number of knocked out BSs when there are L damages in the RAN.

It is assumed that the probability of damages arising in RAN physical links corresponds to the Poisson process [54]. Using the Poisson distribution the probabilities p_L can be expressed over the probability that there are no damages in the RAN p_0. Besides, it was argued in [P6] that there is no practical sense to perform the calculations for the damages number more than 3. Hence, the expression (5.1) takes the following view

$$\begin{aligned} \hat{X}(N,M,L) &= \hat{X}(N,M/L=1)(-\ln p_0)p_0 + \\ &+ \hat{X}(N,M/L=2)\frac{(-\ln p_0)^2 p_0}{2} + \hat{X}(N,M/L=3)\frac{(-\ln p_0)^3 p_0}{6} \cdot \end{aligned} \tag{5.2}$$

Since BS may be considered as knocked out if there are damages in the ring on both side of this BS, the average number of knocked out BSs when a single damage in the RAN $\hat{X}(N,M/L=1)$ is equal to 0.

The rest two coefficients $\hat{X}(N,M/L=2)$, $\hat{X}(N,M/L=3)$ in (5.2) may be calculated as follows

$$\hat{X}(N,M/L=2) = \frac{2(M+2)}{3(N+1)} \qquad \hat{X}(N,M/L=3) = \frac{(M+2)(2N+1)}{(N+1)(N+2)} . \tag{5.3}$$

Since the deduction of the expressions (5.3) is not given in [P6] because of large amount of calculations we present it in the Annex of the thesis.

The expression (5.1) can be written taking into account (5.2), (5.3) and that $\hat{X}(N,M/L=1)=0$ as

$$\hat{X}(N,M,L) = \frac{(M+2)(-\ln p_0)^2 p_0}{3(N+1)} + \frac{(M+2)(2N+1)(-\ln p_0)^3 p_0}{6(N+1)(N+2)} . \tag{5.4}$$

The formula (5.4) is the final expression for calculation of the average number of BSs that are not able to serve mobile terminals as a result of damages in the RAN ring topology.

Ii is interesting to compare the ring topology with other ones from the viewpoint of the reliability parameter. Note that the radial topology is the basic configuration of the RAN physical links of 3G systems [14]. So, in [P6] the expression for calculating the average number of BSs that are not able to serve mobile terminals as a result of damages in the RAN radial topology was obtained

$$\hat{X}(N=0,M,L) = (-\ln p_0)p_0 + (-\ln p_0)^2 p_0 + \frac{(-\ln p_0)^3 p_0}{2}. \tag{5.5}$$

As seen from (5.5) this reliability parameter does not depend on the number of BSs in the RAN at given conditions.

The probability of zero damage (p_0) in the physical links of the RAN may be considered as the availability factor in term of the reliability theory [55]. According to [56], an availability objective more than 99,99% can be considered as acceptable in our case.

Relationships of the average number of knocked out BSs from the rings number N and the number of BSs in each ring M when value p_0 equals 0.99999 is shown in Figure 10.

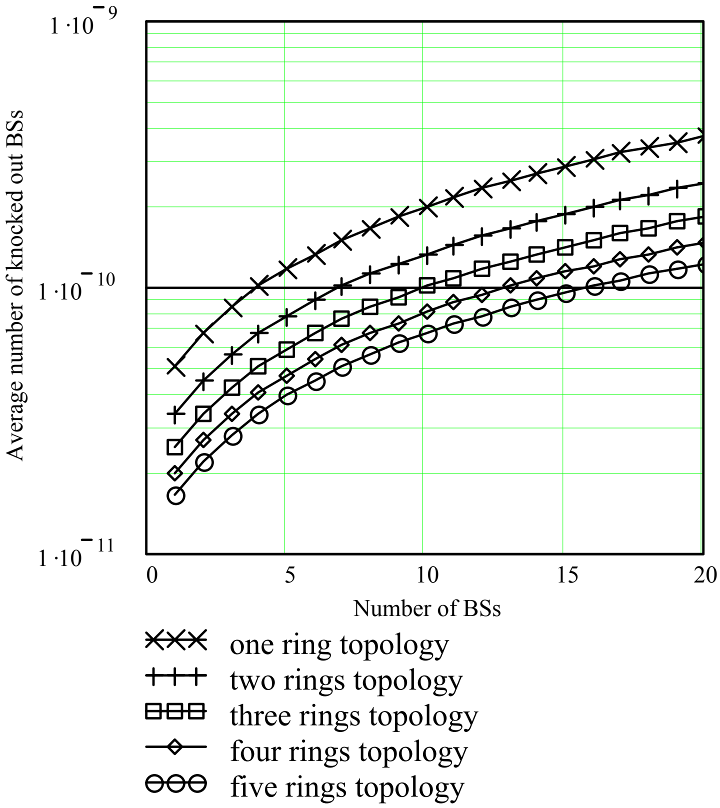

Figure 10. Relationships of the average number of knocked out BSs from the number of rings and the number of BSs in each ring when the availability factor is 0.99999

It is seen from the Figure 10 that value of the average number of knocked out BSs for the ring topologies is order of 10^{-10} when the availability factor equals 0.99999. For comparison, the average number of knocked out BSs for the radial topology is equal to 10^{-5} when the same availability. This value is about five orders larger than one for ring topologies. It means that a radial topology is considerably worse than ring topologies from the viewpoint of reliability.

So, ring topology provides very high reliability for the RAN. For better visualization of this fact let us consider some figures when the availability is very low, for example, when $p_0 = 0.1$. Note that trends of curves shown in Figure 10 do not change when availability value is changed (See Figures 2 and 3 in [P6]). Thus, for one ring in the RAN the average number of knocked out BSs is 1 ($\hat{X}(N,M,L) = 1$) if the number of BSs in the ring is 3. For two rings $\hat{X}(N,M,L) = 1$ if the number of BSs in each ring is 10. For three rings $\hat{X}(N,M,L) = 1$ if the number of BSs in each ring is 7. For four and five rings $\hat{X}(N,M,L) = 1$ if the number of BSs in each ring is equal to 9 and 10 respectively. Therefore, with increasing the ring number the RAN reliability is considerably raised up. Thus, the radial topology is the least reliable topology, the one ring topology is much more reliable than radial topology, and the multi-ring topology (two and more rings) may be considered as the most reliable one.

5.3 Ring topology cost estimation

As it was mentioned in the previous section, expenses for optical fiber physical links are still very essential in the RAN total cost. For this reason, the normalized length of links connecting BSs to each other is considered as a cost quantitative parameter.

The problem of ring topology cost estimating may be formulated as follows. The number of BSs in the RAN, $M \in \{1,2,\ldots\}$; the number of rings in the RAN, $N \in \{1,2,\ldots\}$; coordinates of BSs are given. It is necessary to determine the normalized length of links connecting BSs to each other depending on the number of BSs and rings in the RAN - $L(M,N)$.

The model for ring topology cost estimation is presented in [P6]. The ring topology is organized by type of "petals of flower". For this model the following expression for estimating the cost parameter $L(M,N)$ of the RAN ring topology was obtained in [P6]

$$L(M,N) = \frac{2r(M+N)}{M}, \tag{5.6}$$

where r is the radius of the internal circle of cells.

The expression for calculation of the cost parameter $L(M,N=0)$ of the RAN radial topology has the following view

$$\begin{cases} L(M,N=0) = 2r, \quad M = \overline{1,6} \\ L(M,N=0) = \dfrac{12r + 3{,}46r(M-6)}{M}, \quad M = \overline{7,12} \\ L(M,N=0) = \dfrac{r(4M-15)}{M}, \quad M = \overline{12,18} \end{cases}. \tag{5.7}$$

Relationships of the normalized length of links from the number of BSs and the number of rings in the RAN are presented in Figure 11.

It is seen from Figure 11 that if the number of BSs in the RAN is from 1 to 7 then the cost parameter of the radial topology is more beneficial than one relating to ring topologies. If the number of BSs is 7 and more then ring topologies are preferable to radial from the viewpoint of cost.

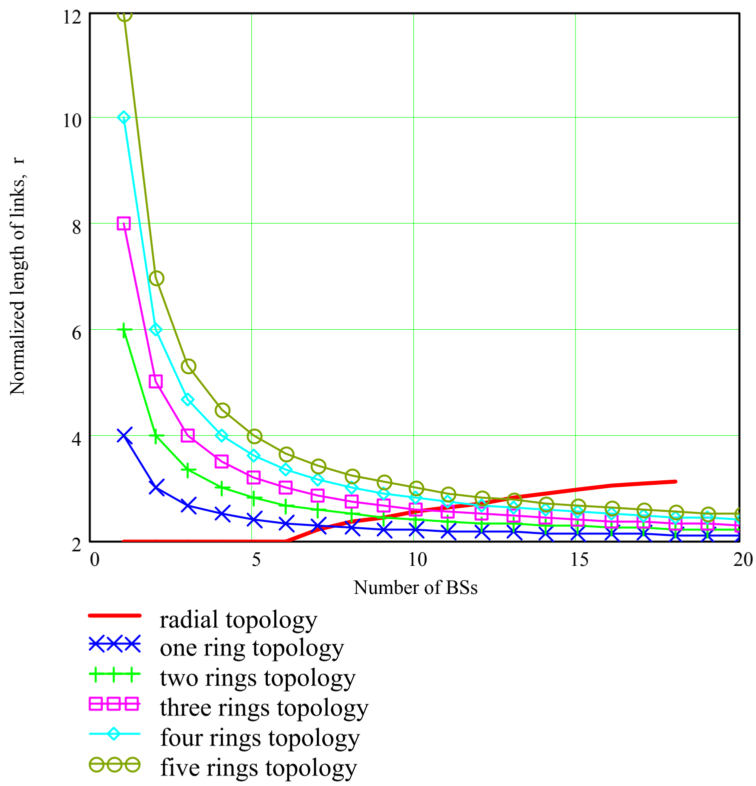

Figure 11. Relationships of the normalized length of links from the number of rings and the number of BSs in the RAN

5.4 Recommendations on using different topologies in the 4G RAN

To give recommendations which of physical link topologies (the radial topology, the one ring topology or the multi-ring topology) should be chosen as the most preferable for 4G RAN deployment it is necessary to consider these topologies from reliability and cost points of view simultaneously. Analyzing both the reliability and cost relationships presented in Figure 10 and Figure 11 together the following conclusions may be formulated.

If the number of BSs in the RAN is 1,2 or 3 then it is worthwhile to apply the radial topology of physical links configuration. However, taking into account very low reliability of the radial topology it may be reasonable to use one ring topology when there are 3 BSs in the RAN. One ring topology is also preferable when there are 4 and 5 BSs in the RAN. If the number of BSs is more than 5 then it is not recommended to apply one ring topology because it has reliability parameters worse than multi-rings topologies. If BS number is from 6 to 10 it is worthwhile to arrange two rings in the RAN. If the number of BSs is from 11 to 16 then three rings topology may be applied. If there is a need for very high reliability then two rings and three rings topologies may be organized when there are 5 BSs and 10 BSs in the RAN respectively.

5.5 Summary

In this chapter, the method for selecting 4G RAN topology of physical link configuration on cost and reliability criteria has been presented. The proposed method allows making quantitative estimations of reliability and cost parameters of different topologies depending on the number of BSs and the number of rings in the RAN. In particular, the parameters of one ring topology, multi-ring topologies (two and more rings), and a radial topology were considered and compared with each other. The results of the research allow making a conclusion that in most cases multi-ring ring topologies are more preferable from the viewpoint of cost and reliability for the 4G RAN than other ones. It should be emphasized that at the final decision about a RAN configuration it is necessary to take into account load parameters and complexity of the network management, as well.

6 THE MINIMUM COST RING CONFIGURATION OF PHYSICAL LINKS BETWEEN BASE STATIONS IN 4G RAN

6.1 Definition of the configuration problem of physical links

If the ring topology (one ring or multi-ring) is chosen for 4G RAN as a result of the previous problem solution then one more problem of effective 4G RAN planning arises. This is the problem of minimum cost configuration of physical links within "a ring". As an example, one of the possible 4G-RAN ring configuration of physical links for the ring # 3 is illustrated in Figure 12. However, it is necessary to determine the minimum-cost configuration for the ring of all possible ring configurations.

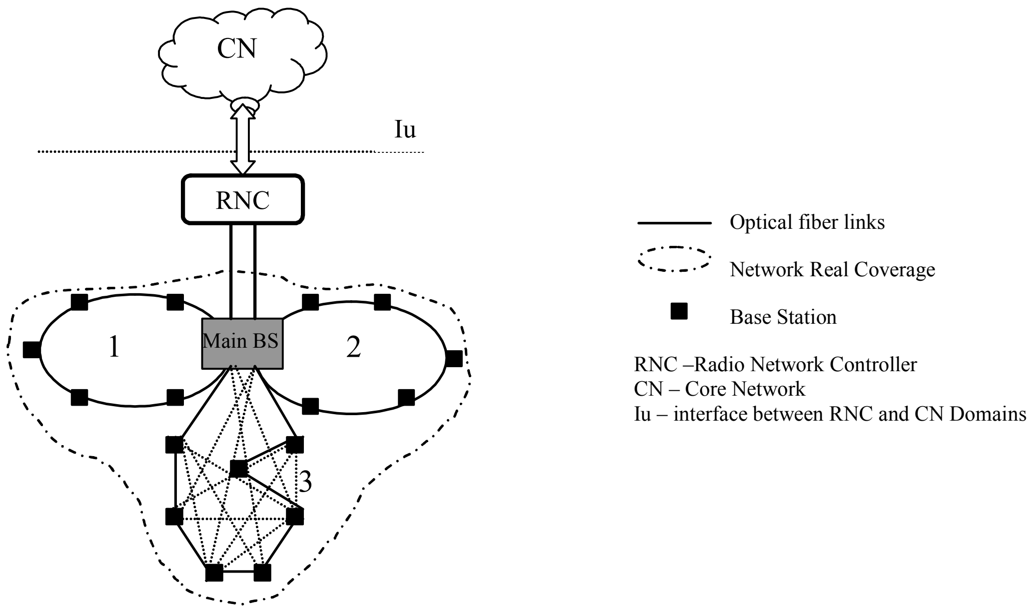

Figure 12. An example of the 4G-RAN ring configuration

The problem of the minimum cost ring configuration of physical links between BSs in the 4G RAN is formulated as follows [P7].

The number of BSs, $M \in \{1,2,\dots\}$ and the cost matrix of all possible links connecting these 4G-RAN nodes with each other are given. It is assumed that BSs are connected to each other by optical fiber links.

It is necessary to find the shortest path (from the viewpoint of the cost optimum criterion) that satisfies the following constraints:

i. It must pass over all BSs.

ii. It must pass over each BS only once.

iii. It must begin and end in the same point corresponding to the "main" BS.

iv. It must have no splits.

Let us number the nodes: $x \in \{1,2,\dots,M\}$.

Then, the problem is to determine

$$\min \sum_{i=1}^{M} C(x_{i-1}, x_i) = \min_{\{x_2,\dots,x_{M+1}\}} J_{M+1}(x_1, x_2, \dots, x_M, x_{M+1}), \qquad (6.1)$$

taking into consideration the above-stated constraints. Here, x is the number of a node, $x \in \{1,2,\ldots,M\}$; the index i denotes the place of a node on the path, i=1,2,...,(M+1); $C(x_{i-1},x_i)$ is the cost of a link between x_{i-1} and x_i nodes respectively; $J_{M+1}(x_1,x_2,\ldots,x_M,x_{M+1})$ is the path cost equal to the total cost of the links in it.

6.2 Approach for the problem solution

The search problem of shortest-path is posed in the graph theory as a combinatorial-optimization problem of high complexity (NP-complete) [57].

Most of the known shortest-path algorithms cannot be applied for the problem solution because of the above-stated constraints [57]. For many of such problems optimal solutions cannot be guaranteed [57]. The inability to differentiate between routes of minimum cost and near-minimum cost is definitely a limitation of use of neural networks for shortest-path searching. It may be used only to perform shortest path computations in applications which do not require an extremely high degree of accuracy [62]. In our study it is proposed to use the dynamic programming principle to solve the problem [58]. Note that the dynamic programming principle has been successfully used for network planning purposes [59,60].

It is necessary to find the shortest path beginning in the node x_1=1, which corresponds to the "main" BS location, and coming to its end in the node x_{M+1}. It is clear, that the node x_{M+1} coincides with the node x_1: $x_1=x_{M+1}=1$.

There are nodes from 2 to M on intervals i=2,3,...,M. The state of the system on i-step is given as $x_i \in \{2,3,\ldots,M;\ i=2,3,\ldots,(M+1)\}$. The form of the objective function for the path with its length of i-links has the following view

$$\underset{i=1,2,\ldots,(M+1)}{J_i}(x_1,x_2,\ldots,x_i) = J(x_i) = \begin{cases} \hat{J}(x_{i-1}) + C(x_i,x_{i-1}), & if \quad x_i \cap \{\hat{x}_i \ldots \hat{x}_{i-1}\} = 0 \\ \infty, \quad if \quad x_i \cap \{\hat{x}_i \ldots \hat{x}_{i-1}\} \neq 0, \end{cases} \qquad (6.2)$$

where $\hat{J}(x_{i-1})$ is the survive path cost equal to the sum of link costs from the node 1 to the node x_{i-1}.

Using the Bellman principle [58] the system of M functional equations (for each value x_i) is formulated as

$$\hat{J}(x_i) = \begin{cases} \min\limits_{x_i \to x_{i-1}} \hat{J}(x_{i-1}) + C(x_i,x_{i-1}), & if \quad x_i \cap \{\hat{x}_i \ldots \hat{x}_{i-1}\} = 0 \\ \infty, \quad if \quad x_i \cap \{\hat{x}_i \ldots \hat{x}_{i-1}\} \neq 0 \quad (*) \end{cases} \qquad (6.3)$$

A solution of the system (6.3) is the minimum-cost path that may be considered as a route for the physical link ring configuration. In cases when it is planned to connect a large amount of nodes (more than 7) in one ring it is worthwhile, just in case, to repeat the above-stated search procedure for the path beginning in other nodes: $x_1 = x_{M+1} = 2,3,\ldots,M$. It will make sure that the variant of configuration determined earlier is right. Emphasis that it has been argued in [P6] that the number of BSs in one ring does not have to exceed 6 or 7.

Note that the restriction (*) allows taking into account the constraint (ii) according to which the path should pass only once over each node. Using such approach it is possible to achieve the following advantages:

- to get the minimum-cost ring path;

- to reduce the huge amount of calculations taking place in such kind of problems up to $\sim M^3$ (M is the number of nodes);
- to take into account all the above-stated constraints;
- to observe the non-ring configurations close to the optimum on intermediate stages of the search procedure.

These advantages make the approach different from other search algorithms [57].

6.3 Case Study

The proper computer software for searching the minimum-cost ring configuration was realized on a basis of the above-mentioned approach. The software calculations were able to produce minimum-cost results in all runs. An example of such kind of calculations is presented below.

There are 7 BSs ($M = 7$) in the 4G RAN. The "main" BSs connected to the dedicated RNC is located in the node # 7. The cost matrix of the links connecting these nodes has the following view

$$\|C\| = \begin{Vmatrix} c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16} & c_{17} \\ c_{21} & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} & c_{27} \\ c_{31} & c_{32} & c_{33} & c_{34} & c_{35} & c_{36} & c_{37} \\ c_{41} & c_{42} & c_{43} & c_{44} & c_{45} & c_{46} & c_{47} \\ c_{51} & c_{52} & c_{53} & c_{54} & c_{55} & c_{56} & c_{57} \\ c_{61} & c_{62} & c_{63} & c_{64} & c_{65} & c_{66} & c_{67} \\ c_{71} & c_{72} & c_{73} & c_{74} & c_{75} & c_{76} & c_{77} \end{Vmatrix} = \begin{Vmatrix} 0 & 0.333 & 0.048 & 0.952 & 0.905 & 0.714 & 0.476 \\ 0.333 & 0 & 0.381 & 1 & 0.857 & 0.286 & 0.429 \\ 0.048 & 0.381 & 0 & 0.524 & 0.095 & 0.762 & 0.571 \\ 0.952 & 1 & 0.524 & 0 & 0.809 & 0.238 & 0.190 \\ 0.905 & 0.857 & 0.095 & 0.809 & 0 & 0.619 & 0.143 \\ 0.714 & 0.286 & 0.762 & 0.238 & 0.619 & 0 & 0.666 \\ 0.476 & 0.429 & 0.571 & 0.190 & 0.143 & 0.666 & 0 \end{Vmatrix} \quad (6.4)$$

It is necessary to find the ring shortest path passing over all BSs only once.

The illustration of the search procedure is shown in Figure 13. Software calculations give the shortest path passing over the following nodes: 7, 5, 3, 1, 2, 6, 4, 7, where node # 7 corresponds to the "main" BS location. If we repeat the above-stated path search procedure starting with other nodes then the minimum-cost path passes by the same route. For instance, starting with the node # 6 the minimum-cost ring path passes over the nodes 4, 7, 5, 3, 1, 2, 6. Therefore, such route of the ring configuration of physical links can be considered as acceptable for the 4G RAN planning.

The cost of the ring shortest path is equal to 1.333. In contrast, the cost of the configuration of radial links from the "main" BS (the node # 7) to other BSs is 2.475. It is of 86 % more than the cost of the ring route.

Let us compare the value of the ring route with the value of the cost bottom boundary of path that satisfies the above-formulated constraints (i) and (ii) only. The cost bottom boundary is calculated as follows

$$C_{opt} \geq \sum_{i=1}^{M} \min_j C(x_i, x_j), \qquad i \neq j, \qquad j = \overline{1, M}, \quad (6.5)$$

where $C(x_i x_j)$ is the cost of the link between x_i and x_j nodes.

The non-ring path satisfying the requirement (6.5) passes over the following nodes: 1, 3, 5, 7 ("main" BS), 4, 6, 2. Frequently, such configurations do not have the practical realization because of very low reliability and routing capabilities. The cost of such non-ring path comes to 1. The difference between the value and the value of the minimum ring route cost is 0.333. In other words, the difference is equal to the cost of the link between the nodes # 1 and # 2. If we add the link to the non-ring path then we get the minimum cost ring configuration obtained earlier. This fact confirms the robustness of the proposed approach, and allows suggesting that it may perform well in considerably larger examples as well.

It should be emphasized that such software calculations were run for different number of BSs in the 4G RAN and in all cases ring configurations close to the bottom boundary of cost (6.5) were obtained.

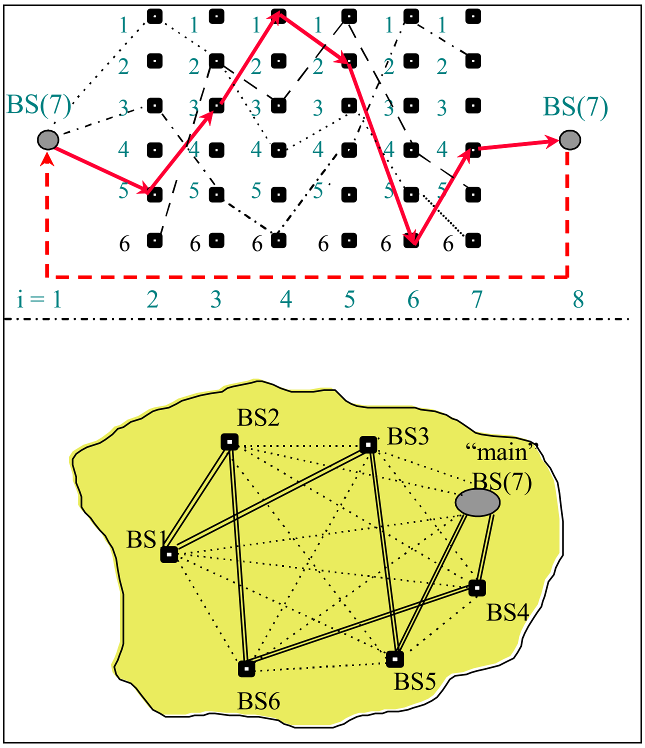

Figure 13. An example of the determination of the ring configuration with the minimum cost for 7 BSs in 4G RAN

6.4 Summary

In this chapter, the configuration problem of physical links between BSs for the 4G RAN ring topology has been considered as a combinatorial-optimization one. The ability of dynamic programming principle to provide an acceptable solution for the problem was demonstrated.

The proposed approach enables getting the minimum cost ring configuration of optical fiber links between BSs in the 4G RAN that allows reducing the expenses on fiber-optic cable and cable installation while a 4G RAN deployment. This approach is characterized by the small amount of calculations (up to $\sim M^3$, where M is the number of BSs in the ring) that is not typically for such kind of problems. One more advantage of the submitted approach from other search algorithms, which can be used for similar problems, is visualization of the state of the system on each step of the search procedure. In other words, the non-ring configurations close to the bottom boundary of cost (5.5) can arise on intermediate stages. It enables considering different variants of topologies of physical links when planning a 4G RAN.

Software calculation results demonstrate the effectiveness of the approach for the minimization of the ring configuration cost. Note that the same approach may be easily adopted for the similar link configuration problems between such elements of the 4G RAN as Radio Network Controllers.

7 CONCLUSION

Wireless network planning is a very complex process, the result of which influences on the success of network operators. A poorly planned network cannot achieve the required Quality of Service. It also involves extra costs and fewer benefits for its network operator. Actually, wireless network planning deals with a large number of different aspects. Since there has been evolution in wireless communications towards the third generation (3G) and fourth generation (4G) wireless systems some Core Network planning aspects for the 3G wireless networks and some Radio Access Network planning aspects for the 4G wireless networks have been considered in this thesis.

In accordance with the main objective of the research some methods for effective network planning were developed to get some technical and economical benefits when creating and exploiting the 3G/4G wireless networks.

The main contributions of the thesis to 3G CN planning are as follows:

- The method for estimating the number of prospective 3G users in a region. If the number of population and the penetration level of 2G services in the region are known then the method allows estimating a demand for 3G services in the region before 3G equipment deployment. The method includes the novel approach for estimating the Pareto parameter based on the use of Lorenz curves. The obtained estimations are used as input data for the estimation problem of 3G data traffic characteristics.
- The method for estimating the main characteristics of data traffic generated by 3G users. The method allows combining separated statistical data concerning 3G service providing into the unified probabilistic model of events relating to calls of 3G services. The model is based on both decomposition of a set of 3G services into three subsets in accordance with an average amount of transferred data per transaction and distribution of 3G users into three subgroups in accordance with their demand for 3G services from the different subsets. It enables segregating all traffic into several segments. On a basis of the probabilistic model the following main parameters of 3G data traffic are estimated:
    - the average duration of 3G calls for each subset of services,
    - the rate of transactions (per user) in busy hour for each traffic segment,
    - the traffic intensity (per user) generated by 3G calls for each traffic segment,
    - total traffic intensity on nodes of the PS CN domain and others.

  The estimations of the traffic parameters are used to adjust with them the performance measures of IP Multimedia Core Network subsystem (IM-subsystem) nodes.
- The method for performance evaluation of IM-subsystem nodes taking into account the influence of self-similar input. The method is based on the use of FBM/D/1/W queueing system, where FBM is Fractional Brownian Motion. The following main performance measures are estimated with help of the method:
    - upper and lower bounds for the node service rate,
    - upper and lower bounds for the average queue length in the node buffer,
    - upper and lower bounds for the average service time of information units in the node buffer,
    - the node utilization.

The main contributions of the thesis to 4G RAN planning are as follows:

- The method for a quantitative estimation of reliability and cost parameters of the 4G RAN ring topology since this topology may be applied in the 4G RAN in contrast to RANs of current wireless systems. The method allows comparing one ring topology, multi-ring topologies (two and more rings) and radial topology with each other and choosing the most preferable one on cost and reliability criteria for 4G RAN deployment depending on the number of base stations in the RAN.
- The method for the minimum cost ring configuration of physical links between base stations in the 4G RAN if the ring topology is chosen for 4G RAN deployment.

The results of the first part of the research allow establishing requirements, such as throughput of channels, performance of nodes and routers, both to the equipment of internal IP-networks in the 3G wireless systems and to the gateways providing the interaction between 3G systems and external networks.

Estimations of data traffic parameters and performance measures of IP-network equipment obtained with help of the proposed methods allow making easier both the preparation of business-plans for development of wide range of 3G services and the preparation of system projects for deployment of PS CN domain and, afterwards, IM-subsystem.

Results of the second part of the research enable minimizing expenses with the 4G RAN deployment taking into account reliability requirements. Besides, some recommendations about using different topologies of physical link configuration in the RAN for mobile communication systems beyond IMT-2000 were given.

The developed methods for efficient 3G/4G network planning are considered as a basis for software package designing in order to automate and to optimize the planning process.

Further research is to extend the obtained results concerning 3G Core Network Planning for 4G wireless networks and fixed NGN from the viewpoint of user traffic estimation, performance evaluation, network architecture. We are going to estimate signaling traffic using Signaling System No. 7 (SS7) and Session Initiation Protocol (SIP) for 4G wireless networks and fixed NGN. The signaling traffic is generated by support procedures of Customized Application for Mobile Network Enhanced Logic (CAMEL) and service mobility based on a Mobile Application Part (MAP) protocol and Mobile IPv4/IPv6. Such estimations of the signal traffic may be useful for the realization of the Virtual Home Environment (VHE) conception. Moreover, it is planned to analyze some more RAN physical link configuration solutions from viewpoint cost and reliability. Finally, it would be interesting to develop the method based on Net Present Value (NPV) modeling for estimating the effectiveness of investment in different projects concerned with the deployment of 3G/4G wireless systems and the development of 3G/4G services.

8 SUMMARY OF PUBLICATIONS

8.1 Overview of publications

In accordance with the structure of the thesis all publications can be divided into two groups. The first group contains the publication concerning Core Network planning aspects for the 3G wireless systems [P1]-[P5]. The second group includes the publication about Radio Access Network planning aspects for the 4G wireless system [P6]-[P8].

P1. Krendzel A., Koucheryavy Y., Derzhavina V., Harju J. "*Method for Estimating the Number of Potential 3G Users*", in proceedings of 10-th Open European Summer School on the Advances in Fixed and Mobile Networks (EUNICE – 2004), Tampere, Finland, June 14-16, 2004.

Description The prediction problem of a demand for 3G services in a region before installation of 3G wireless systems is considered in [P1]. In particular, the method for preliminary estimating the number of potential users of 3G services is proposed. It is based on using the Pareto law, the Lorenz curves, the Gini coefficient and some economic rules. The relationship between the relative number of 3G users and penetration level of 2G services is derived. Besides, the approach for estimating the Pareto parameter is presented in [P1]. The obtained results may be considered as initial data for other problems of network planning, in particular, for the estimation problem of 3G data traffic parameters.

P2. Krendzel A., Koucheryavy Y., Lopatin S., Harju J. "*Method for Estimating Parameters of 3G Data Traffic*", in proceedings of the IEEE International Conference on Communications (ICC-2004), Paris, France, June 20 – 24, 2004.

Description The prediction problem of 3G data traffic characteristics is considered in [P2]. The method for estimating parameters of 3G data traffic is based on probabilistic model of events initiated by 3G calls. This model is composed in conformity with rules for decomposition of 3G services into three subsets taking into account the average amount of transferred data during a transaction and the approximation of users' distribution for each subset of 3G services. Users' distribution for each subset of 3G services is approximated in accordance with the Pareto law.

P3. Krendzel A., Derzhavina V., Lopatin S. "*Method for Estimating Parameters of NGN traffic*", in proceedings of the International Conference on Next Generation Teletraffic and Wired/Wireless Advanced Networking (NEW2AN), pp. 50-54, St.-Petersburg, Russia, February 02-06, 2004.

Description The method presented in previous publication is developed in [P3] for the case when it is necessary to estimate the main characteristics for both symmetrical and asymmetrical traffic separately depending on a type of 3G services generating the traffic. In particular, it is proposed to use other principles for the service decomposition in this case. In the first step of the decomposition, a set of services is divided into two subsets. The first subset comprises of services concerning the real-time establishment of connectivity between endpoints. It is characterized by the transfer of the symmetrical traffic and the strict control of QoS. The second subset deals with services that generate the asymmetrical traffic. In the second step of the decomposition, each of the above-mentioned subsets of services is divided into three classes in accordance with features of the generated traffic intensity. The division of the first subset of services creating symmetrical traffic into three classes is fulfilled taking into account the flow rate initiated by users. The division of the second subset of services initiating the asymmetrical traffic into three classes is fulfilled taking into account an average amount of information initiated by users. The principles for distribution of users into three subgroups are the same as in [P2]. The main parameters for symmetrical and asymmetrical traffic arriving on network equipment are estimated separately using common probabilistic model. The probabilistic model includes three entire groups of statistically independent events and allows segregating 18 segments from the common data traffic.

P4. Koucheryavy Y., Krendzel A., Lopatin S., Harju J. "*Performance Estimation of UMTS Release 5 IM-Subsystem Elements*", in proceedings of the 4-th IEEE Conference on Mobile and Wireless Communications Network (MWCN 2002), Stockholm, Sweden, September 9-11, 2002.

Description In [P4] the estimation problem of performance measures of IP Multimedia Core Network subsystem nodes (IM – subsystem) in UMTS Release 5 is considered. In particular, the method for performance evaluating of Gateway GPRS Support Node (GGSN) is proposed on a basis of FBM/D/1/W queueing system. The method may be applied for evaluation of the lower and upper bounds of the main performance measures of IM-subsystem nodes taking into consideration the self-similar nature of the multiservice traffic. Besides, the vision of UMTS Rel'5 architecture based on configurations from 3GPP TS 23.002 is presented in [P4].

P5. Koucheryavy Y., Krendzel A. "*Analysis of Web Traffic and Users' Behavior Modeling During Busy Hour*", in proceedings of the 6-th Open European Summer School (EUNICE – 2000), University of Twente, Enschede, the Netherlands, September 13-15, 2000.

Description This publication deals with results of statistical analysis of WWW server under heavy load. Behavior of users, in particular, dynamic of user requests towards the web server during the busy hour is analyzed in [P5]. The study of the obtained data concerning the observed web traffic demonstrates that the particular traffic has self-similar nature. In particular, for the collected statistical data the Hurst parameter characterizing a measure of burstiness for self-similar series is approximately equal to 0.85.

P6. Krendzel A., Koucheryavy Y., Harju J. "*Cost and Reliability Estimation of Radio Access Network Structures for the 4G Systems*", in proceedings of the IEEE 58-th Vehicular Technology Conference (VTC 2003-Fall), the Symposium on Wireless Communications: 3G and Beyond, Orlando, USA, October 6-9, 2003.

Description The publication [P6] is devoted to architecture aspects of Radio Access Network physical link configuration for the 4G wireless networks. Firstly, reliability parameters for topologies consisting of one, two, three and more rings and radial topology are estimated. Then, the cost parameters of these structures are estimated. After that, the topologies are compared with each other taking into consideration both cost and reliability. Principles of probability theory are used for such approach. Finally, some recommendations on applying different physical link topology to the RAN of the 4G wireless systems are given.

P7. Krendzel A., Koucheryavy Y., Harju J. "*Radio Access Network Topology Planning for the 4G Networks*", in proceedings of the 5th European Wireless Conference Mobile and Wireless Systems beyond 3G (EW-2004), Barcelona, Spain, February 24 – 27, 2004.

Description The configuration problem of physical links between base stations within the ring topology from the viewpoint of minimum 4G RAN deployment cost is considered in [P7]. It is formulated as a combinatorial-optimization problem. The method on a basis of dynamic programming principles is applied for its solution. Software calculation results demonstrate the capability of the method to determine the minimum cost ring configuration.

P8. Krendzel A., Koucheryavy Y. "*Remote Units Location Problem in Multiservice Access Network*", in proceedings of International Conference on Telecommunications (ICT-2002), Beijing, China, pp. 496-500, June 23-26, 2002.

Description The location problem of remote units when planning a new transport multiservice access network is considered in this publication. Resources of the universal access network are used by set of switched networks such as the 3G and 4G wireless networks, CSPDN, Internet, PSTN and others. The method for determining coordinates of remote units to give the minimum access network deployment cost is proposed in [P8]. It is based on the use of sphere packing principles. It is possible to locate in the obtained points as the equipment for wireless access as the equipment for wired access.

8.2 Author's Contribution to the Publications

The contribution of the author is significant in all of the publications. In the publications [P1], [P2], [P3], [P6], [P7], [P8] he has been the main author, and in the papers [P4], [P5] he has contributed to an essential part of the content of the publications. The author has developed the theoretical framework, performed the experiments and prepared the manuscript. The main contribution of author in each publication is as follows.

In [P1] the author proposed to apply relationships between the Pareto law, the Lorenz curves, the Gini coefficient, Gross Domestic Product, income distribution between population, and infocommunication density for estimating a demand for 3G services in a region where it is planned to deploy 3G network equipment.

In [P2] he proposed to make decomposition of 3G services into some subsets in accordance with an average amount of information transferred during transactions and distribution of 3G prospective users into some subgroups in accordance with their demand for 3G services from the different subsets that allows forming the probabilistic model of events initiated by 3G calls, and estimating, after that, the main parameters of 3G data traffic.

In [P3] the author developed the method presented in [P2] for the case when it is necessary to estimate parameters for symmetrical and asymmetrical traffic separately depending on a type of services generating the traffic. In particular, other rules for decomposition of services were introduced.

In [P4] the author analyzed the probabilistic and time characteristics of Gateway GPRS Support Node (GGSN) on a base of FBM/D/1/W queueing system, where FBM is Fractional Brownian Motion.

In [P5] the author examined the obtained statistical data concerning the web traffic for self-similarity. In particular, the value of the Hurst parameter for the observed fragment of the web traffic was found.

In [P6] the author introduced quantitative parameters for estimating cost and reliability of the new ring 4G RAN topology and proposed to analyze the parameters simultaneously. As results, some recommendations how to choose different types of physical link topologies depending on the number of BSs in 4G RAN were developed.

The main contribution of the author in [P7] was the use of the dynamic programming principle for the problem solution of the physical link ring configuration between BSs from the viewpoint of minimum 4G RAN deployment cost.

In [P8] the author formulated the location problem of remote units for multiservice access network and considered the problem solution using the sphere packing principles.

ANNEX

Calculating the average number of knocked out Base Stations when there are two or three damages in the 4G RAN ring topology

In the Annex the expressions for calculating the average number of knocked out Base Stations (BSs) when there are two or three damages in the 4G RAN ring topology are derived. The expressions are used in the fifth chapter as coefficients $\hat{X}(N, M / L = 2)$ and $\hat{X}(N, M / L = 3)$ respectively in the formula (5.2) for estimating the average number of knocked out BSs depending on the number of rings in the RAN, the number of BSs in the ring, the damages number in the RAN $\hat{X}(N, M, L)$.

The deduction of the expressions is based on use of probability theory principles.

The common expression for the average number of knocked out BSs when there are 2 damages in the RAN $\hat{X}(N, M / L = 2)$ has the following view

$$\hat{X}(N, M / L = 2) = \hat{X}(N_{1+1}, M) p(N_{1+1}) + \hat{X}(N_{2+0}, M) p(N_{2+0}) \tag{A.1}$$

where $\hat{X}(N_{1+1}, M)$ is the average number of knocked out BSs when there are two damages in the RAN taking place in different rings, $p(N_{1+1})$ is the probability that two damages occur in different rings, $\hat{X}(N_{2+0}, M)$ is the average number of knocked out BSs when there are two damages taking place in one ring, $p(N_{2+0})$ is the probability that two damages occur in one ring, M is the number of BSs in a ring, $M \in \{1,2,\ldots\}$.

It is clear that $\hat{X}(N_{1+1}, M) = 0$. The parameter $\hat{X}(N_{2+0}, M)$ may by calculated using the following formula

$$\hat{X}(N_{2+0}, M) = \sum_{n=1}^{M} x_n p_n, \tag{A.2}$$

where $x_n = M - n + 1$, $x_n \in \{M, M-1, \ldots, 1\}$, $p_n = k/B$, $k = n$, $B = C_{M+1}^2$, x_n is the number of knocked out BSs, B is the entire group of events of knocked out BSs: (1,2,…,M) BSs, k is the number of different variants (multiple factor) of knocked out x_n BSs, p_n is the probability of knocked out x_n BSs, C_{M+1}^2 is the number of combination of $M+1$ BSs taken 2 at a time.

Taking into consideration x_n and p_n the expression for $\hat{X}(N_{2+0}, M)$ may be rewritten as

$$\hat{X}(N_{2+0}, M) = \frac{1}{C_{M+1}^2} \sum_{n=1}^{M} n(M - n + 1) = \frac{M+2}{3}. \tag{A.3}$$

Now it is necessary to derive expressions for probabilities $p(N_{1+1})$ and $p(N_{2+0})$. Let's denote appearance of two damages in the RAN by the event $N_{L=2}$, appearance of two damages in one ring by the event N_{2+0}, appearance of two damages in two different rings by the event N_{1+1}. Then, the probability of two damages in the RAN is $p(N_{L=2}) = p(N_{2+0}) + p(N_{1+1})$. The total number of outcomes of appearance of two damages in the RAN ($B_{L=2}$) is $B_{L=2} = B_{2+0} + B_{1+1} = N + C_N^2 = N(N+1)/2$, where B_{2+0} is the number of outcomes that are favorable for the event N_{2+0}, B_{1+1} is the number of outcomes that are favorable for the event N_{1+1}, C_N^2 is the number of combination of N rings taken 2 at a time. Hence,

$$p(N_{2+0}) = \frac{2}{N+1} \qquad p(N_{1+1}) = \frac{N-1}{N+1}. \tag{A.4}$$

Taking into account (A.3), (A.4) and that $\hat{X}(N_{1+1}, M) = 0$ the expression for calculating the coefficient $\hat{X}(N, M / L = 2)$ has the following view

$$\hat{X}(N, M / L = 2) = \hat{X}(N_{2+0}, M) p(N_{2+0}) = \frac{2(M+2)}{3(N+1)}. \tag{A.5}$$

The average number of knocked out BSs when there are three damages in the RAN is determined as

$$\hat{X}(N, M / L = 3) = \hat{X}(N_{1+1+1}, M) p(N_{1+1+1}) + \hat{X}(N_{3+0}, M) p(N_{3+0}) + \hat{X}(N_{2+1}, M) p(N_{2+1}) \tag{A.6}$$

where $\hat{X}(N_{1+1+1}, M)$ is the average number of knocked out BSs when there are three damages in the RAN taking place in different rings, $p(N_{1+1+1})$ is the probability that three damages occur in different rings, $\hat{X}(N_{3+0}, M)$ is the average number of knocked out BSs when there are three damages taking place in one ring, $p(N_{3+0})$ is the probability that three damages occur in one ring, $\hat{X}(N_{2+1}, M)$ is the average number of knocked out BSs when there are two damages in one ring and one damage in either of the rest two rings, $p(N_{2+1})$ is the probability that two damages occur in one ring and one damage occurs in either of the rest two rings.

It is obvious that the average number of knocked out BSs when there are three damages in the RAN taking place in different rings is equal to zero $\hat{X}(N_{1+1+1}, M) = 0$

The parameter $\hat{X}(N_{3+0}, M)$ is determined as follows

$$\hat{X}(N_{3+0}, M) = \sum_{n=1}^{M-1} x_n p_n, \tag{A.7}$$

where $x_n = M - n + 1$, $x_n \in \{M, M-1, ..., 2\}$, $p_n = k/B$, $k = n(M-n)$, $B = C_{M+1}^3$, C_{M+1}^3 is the number of combination of $M+1$ BSs taken 3 at a time.

By substituting x_n and p_n into (A.7) we get the following expression

$$\hat{X}(N_{3+0}, M) = \frac{1}{C_{M+1}^3} \sum_{n=1}^{M-1} (M-n+1)(M-n)n = \frac{M+2}{2}. \tag{A.8}$$

It is clear that the expression for calculation $\hat{X}(N_{2+1}, M)$ is the same as the expression (A.3) for $\hat{X}(N_{2+0}, M)$

$$\hat{X}(N_{2+1}, M) = \frac{M+1}{3}. \tag{A.9}$$

Next step of the procedure is to calculate the probabilities for $p(N_{1+1+1})$, $p(N_{2+1})$, $p(N_{3+0})$. Let's denote appearance of three damages in the RAN by the event $N_{L=3}$, appearance of three damages in one ring by the event N_{3+0}, appearance of three damages in three different rings by the event N_{1+1+1}, appearance of two damages in one ring and one damage in either of the rest two rings by the event N_{2+1}, Then, the probability of appearance of three damages in the RAN is $p(N_{L=3}) = p(N_{3+0}) + p(N_{1+1+1}) + p(N_{2+1})$. The total number of outcomes of appearance of three damages in the RAN ($B_{L=3}$) is $B_{L=3} = B_{3+0} + B_{1+1+1} + B_{2+1} = N + C_N^3 + A_N^2 = N(N+1)/2$, where B_{3+0} is the number of outcomes that are favorable for the event N_{3+0}, B_{1+1+1} is the number of outcomes that are favorable for the event N_{1+1+1}, B_{2+1} is the number of outcomes that are favorable for the event N_{2+1}, C_N^3 is the number of combination of N rings taken 3 at a time, A_N^2 is the number of permutation of N rings taken 2 at a time. Therefore,

$$P(N_{1+1+1}) = \frac{(N-1)(N-2)}{(N+1)(N+2)}, \quad P(N_{2+1}) = \frac{6(N-1)}{(N+1)(N+2)}, \quad P(N_{3+0}) = \frac{6}{(N+1)(N+2)}. \quad \text{(A.10)}$$

Taking into account (A.8), (A.9), (A.10) and that $\hat{X}(N_{1+1+1}, M) = 0$ the expression (A.6) takes the following view

$$\hat{X}(N, M / L = 3) = \hat{X}(N_{3+0}, M)\,p(N_{3+0}) + \hat{X}(N_{2+1}, M)\,p(N_{2+1}) = \frac{(M+2)(2N+1)}{(N+1)(N+2)}. \quad \text{(A.11)}$$

Thus, the expressions (A.5) and (A.11) may be used for calculating the average number of knocked out Base Stations when there are two or three damages in the 4G RAN respectively.

REFERENCES

[1] H. Kaaranen, A. Ahtiainen, L. Laitinen, S. Naghian, V. Niemi, “UMTS networks. Architecture, mobility and services,” John Wiley & Sons, 2001.

[2] M. C. Bale, “Voice and Internet multimedia in UMTS networks,” BT Technology Journal, vol. 19, №.1, 2001.

[3] Y.Yamao, H. Suda, N.Umeda, N. Nakajima, “Radio access network design concept for the fourth generation mobile communication system,” Proc. VTC2000-Spring, vol. 3, pp. 2285-2289, 2000.

[4] “The UMTS 3G Market Forecasts – Post September 11, 2001,” Report # 18 from the UMTS Forum, February 2002.

[5] M. Jiang, M. Nicolic, S. Hardy, L. Trajkovic, “Impact of self-similarity on wireless data network performance,” ICC 2001, USA, June 2001.

[6] K. Park, W. Willenger, “Self-Similar network traffic and performance evaluation,” John Wiley & Sons, 2000.

[7] R. Kalden, S. Ibrahim, “Searching for self-similarity in GPRS”, Proceedings of the 5th anuual Passive & Active Measurement Workshop (PAM2004), Antibes Juan-les-Pins, France, April 19-20, 2004

[8] V. Paxson, S.Floyd, “Wide area traffic: the failure of Poisson modeling,” Proceedings of ACM SIGCOMM’94, 1994.

[9] H. Chebbo, M. Wilson, “Traffic and load modelling of an IP mobile network,” Proceedings of the 4th International Conference on 3G Mobile Communication Technologies (Conf. Publ. No. 494), pp. 423 – 427, June 25-27, 2003.

[10] A. Silva, G. Mateus, “Performance analysis for data service in third generation mobile telecommunication networks”, Proceedings of the 35th Annual Simulation Symposium, pp. 227-234, April 14-18, 2002.

[11] A.R.S. Bahai, H. Aghvami, “Network planning and optimization in the third generation wireless networks,” Proceedings of the 1st International Conference on 3G Mobile Communication Technologies (Conf. Publ. No. 471), pp. 441 – 445, March 27-29, 2000.

[12] J. Harmatos, “Planning of UMTS core networks,” Proceedings of the 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Volume 2, pp. 740-744, September 15-18, 2002.

[13] T. Otsu, Y. Okajima, N. Umeda, Y. Yamao, “Network architecture for mobile communications systems beyond IMT-2000,” IEEE Personal Commun. Mag., vol. 8, no. 5, pp. 31-37, October 2001.

[14] T. Otsu, N. Umeda, Y. Yamao, “System architecture for mobile communications systems beyond IMT-2000,” GLOBECOM'01, vol. 1, pp. 538 –542, 2001.

[15] Y. Aburakawa, T. Otsu, Y. Ymao, “Fiber and free-space hybrid optical networking for new generation mobile radio access network,” the 5th international symposium on wireless personal multimedia communications, vol. 2, pp. 586 –590, 2002.

[16] V. Pareto, “Le cours d’economie politique,” Macmillann, London, 1897.

[17] J. A. Schumpeter, “Vilfredo Pareto: in ten great economist from Marks to Keynes,” OUP, 1951.

[18] M. Allais, “Pareto, Vilfredo: contributions to economics,” International Encyclopedia of the Social Sciences, Macmillan and Free Press, vol.11, 1968.

[19] M. O. Lorenz, “Methods for measuring the concentration of wealth,” Amer. Stat. Assoc., no. 9, pp. 209-219, 1905.

[20] C. Gini, “Indici di concentrazione e di dipendenza,” Biblioteca delli’eccomista, no.20, 1922.

[21] P. M. Dixon, T. Mitchell-Olds, R. Woodley, “Bootstapping the Gini coefficient of inequality,” Ecology 68, pp. 1548-1551, 1987.

[22] C. Dagum. “The generation and distribution of income, the Lorenz curve and the Gini ratio,” Econ. Appl. No. 33, 327-367, 1980.

[23] R.G. Addie, M. Zukerman, T.D. Neame, “Broadband traffic modeling: simple solutions to hard problems,” IEEE Communications Magazine, pp. 88-95, August 1998.

[24] I. Norros, “On the use of fractional Brownian motion in the theory of connectionless networks,” IEEE Journal on Selected Areas in Communications, vol. 13, № 6, pp. 953-962, August 1995.

[25] I. Norros, “A storage model with self-similar input,” Queueing Systems, vol. 16, pp. 387-396, 1994.

[26] L.E. Varakin, “The Pareto law and the rule 20/80: distribution of incomes and telecommunication services,” MAC proceedings, pp.3-10, no. 1, 1997.

[27] L.E. Varakin, “Economics, telecommunications and development of the society: macroeconomic mechanisms of telecommunications development,” Electrosvyaz Journal, no. 1, 1994.

[28] A. Jipp, “Wealth of nations and telephone density,” Telecommunication Journal, No.6, 1963.

[29] S. Kotz, N.L. Johnson, C.B. Read, “Encyclopedia of statistical science,” New York, Wiley, 1983.

[30] M. Kendall, A. Stuart, “The advanced theory of statistics. Distribution Theory,” London, C. Griffin & Co., 1962.

[31] P. Derusso, Roy R., Close C., A. Desrochers, “State variables for engineers”, John Wiley & Sons, 1998.

[32] C.V. Brown, P.M. Jackson, “Public sector economics”, Oxford, Blackwell, 1990.

[33] http://www.cdmatech.com/resources/glossary_full.jsp

[34] http://www.cdg.org/resources/Guest_column.asp?columnid=9

[35] “Support of third generation services using UMTS in a converging network environment,” Report # 14 from the UMTS Forum, February 2002.

[36] 3GPP TS 123 060 V.3.5.0 (2000-10), “Technical specification digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); General Packet Radio Service (GPRS); Service description; Stage 2,” 3GPP TS 23.060 version 3.5.0, Release 1999.

[37] 3GPP TR 23.873 V.4.0.0 (2001-03) “3rd Generation Partnership Project (3GPP); technical specification group services and systems aspects; feasible study for transport and control separation in the PS CN domain (Release 4)”.

[38] 3GPP TS 23.002 V.5.5.0 (2002-01), “3rd Generation Partnership Project (3GPP); technical specification group services and systems aspects; network architecture (Release 5)”.

[39] “Naming, addressing & identification issues for UMTS,” Report # 12 from the UMTS Forum, February 2001.

[40] W. Willinger, M.S. Taqqu, W.E. Leland, D.V. Wilson, “Self-similarity in high-speed packet traffic: analysis and modeling of Ethernet traffic measurements,” *Statistical Science*, vol. 10, no. 1, pp. 67-85, 1995.

[41] J.B. Roberts, “Traffic theory and the Internet,” IEEE Communications Magazine, pp. 94-99, January 2001.

[42] F. Geerts, C. Blondia, “Superposition of Markov sources and long range dependence”, The 4th International Conference on Broadband Communications (BC '98), pp. 550-562, 1998.

[43] W. Willinger, M.S. Taqqu, R. Sherman, D.V. Wilson, “Self-similarity through high-variability: statistical analisis of Ethernet LAN traffic at the source level,” IEEE/ACM Transactions on networking, vol. 5, no. 1, pp. 71-86, February 1997.

[44] A. Adas, “Traffic Models in Broadband Networks,” IEEE Communicatios Magazine, pp. 82-89, July 1997.

[45] W. Willinger, D. Wilson, M. Taqqu, “Self-similar traffic modeling for high-speed networks,” *ConneXions*, November 1994.

[46] W. Stallings, “High speed networks. TCP/IP and ATM design principles,” Upper Saddles River, NJ: Prentice-Hall, p. 576, 1998.

[47] L. Kleinrock, “Queueing systems. Computer applications,” John Wiley & Sons, vol.2, 1976.

[48] I. Karatsas, S. Shreve, “Brownian motion and stochastic calculus,” New York: Springer-Verlag, 2nd ed., 1997.

[49] A. Papoulis, “Probability, random variables, and stochastic processes,” McGraw-Hill, New York, 2nd ed., 1984.

[50] L. Kleinrock, “Queueing systems. Theory”, John Wiley & Sons, vol. 1, 1975.

[51] M.E. Crovella, A. Bestavros, “Self-similarity in word wide web traffic: evidence and possible causes,” IEEE/ACM Transactions on networking, vol.5, no.6, pp. 835-846, December 1996.

[52] I. Norros, A. Simonian, J. Virtamo, "The Benes method – a unified approach to ATM FIFO queueing,” Proceedings of St. Petersburg International Teletraffic Seminar, St. Petersburg, pp. 431-449, June 1995.

[53] A. Szlovencsak, I. Godor, J. Harmatos, T. Cinkler, “Planning reliable UMTS terrestrial access networks,” IEEE Commun. Mag., vol. 40, no. 1, pp.67-72, Jan. 2002.

[54] J.-B. Gravereaux, J. Ledoux, “Poisson approximation for some point processes in reliability”, Advances in Applied Probability, 36 (2004), no. 2, 455–470.

[55] Telephone network and ISDN, Quality of service, network management and traffic engineering, CCITT, Blue Book, vol. II, Fascicle II.3, pp. 317-365, Geneva, 1989.

[56] J.L. Spencer, “Availability objectives for fiber in the loop systems,” GLOBECOM'91, vol.1, pp. 572 –577, December 1991.

[57] N. Christofides. “Graph theory. An algorithmic approach”. Academic press, 1975.

[58] G. Korn, T. Korn. “Mathematical handbook for scientists and engineers. Definitions, theorems and formulas for reference and review,” McGraw-Hill Book Company, New York, San Francisco, Toronto, London, Sydney, 1968.

[59] Y. Rapp. “The use of dynamic programming for planning of extensions of conduits and main cable networks in a local exchange area”. Ericsson Rev. No.46:3, pp.102-112, 1969.

[60] Y. Rapp, K. Moum. “The application of dynamic programming for planning of main cable networks”. Ericsson Rev. No.50:3, pp.90-100, 1973.

[61] J. Harmatos, A. Szentesi, I. Godor. “Planning of tree-topology UMTS terrestrial access networks”. PIRMS 2000, London, September, 2000.

[62] T. Haines, J.V. Medanic, “A neural network shortest path algorithm,” Proceedings of the 1994 IEEE International Symposium on Intelligent Control, Columbus, Ohio, USA, pp. 382-387, August 16-18, 1994.

Publication 1

Krendzel A., Koucheryavy Y., Derzhavina V., Harju J.

“*Method for Estimating the Number of Potential 3G Users*

Publication 2

Krendzel A., Koucheryavy Y., Lopatin S., Harju J.

"*Method for Estimating Parameters of 3G Data Traffic*"

Publication 3

Krendzel A., Derzhavina V., Lopatin S.

"*Method for Estimating Parameters of NGN traffic*"

Publication 4

Koucheryavy Y., Krendzel A., Lopatin S., Harju J.

"*Performance Estimation of UMTS Release 5 IM-Subsystem Elements*"

Publication 5

Koucheryavy Y., Krendzel A.

"*Analysis of Web Traffic and Users' Behavior Modeling During Busy Hour*"

Publication 6

Krendzel A., Koucheryavy Y., Harju J.

"*Cost and Reliability Estimation of Radio Access Network Structures for the 4G Systems*"

Publication 7

Krendzel A., Koucheryavy Y., Harju J.

"*Radio Access Network Topology Planning for the 4G Networks*"

Publication 8

Krendzel A., Koucheryavy Y.

"*Remote Units Location Problem in Multiservice Access Network*"

Tampereen teknillinen yliopisto
PL 527
33101 Tampere

Tampere University of Technology
P.O. Box 527
FIN-33101 Tampere, Finland