

Mykhailo Ilchenko
Leonid Uryvsky
Larysa Globa *Editors*

Advances in Information and Communication Technologies

Processing and Control in Information
and Communication Systems

Lecture Notes in Electrical Engineering

Volume 560

Series Editors

Leopoldo Angrisani, Department of Electrical and Information Technologies Engineering, University of Napoli Federico II, Napoli, Italy

Marco Arteaga, Departament de Control y Robótica, Universidad Nacional Autónoma de México, Coyoacán, Mexico

Bijaya Ketan Panigrahi, Electrical Engineering, Indian Institute of Technology Delhi, New Delhi, Delhi, India

Samarjit Chakraborty, Fakultät für Elektrotechnik und Informationstechnik, TU München, München, Germany

Jiming Chen, Zhejiang University, Hangzhou, Zhejiang, China

Shanben Chen, Materials Science & Engineering, Shanghai Jiao Tong University, Shanghai, China

Tan Kay Chen, Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore

Rüdiger Dillmann, Humanoids and Intelligent Systems Lab, Karlsruhe Institute for Technology, Karlsruhe, Baden-Württemberg, Germany

Haibin Duan, Beijing University of Aeronautics and Astronautics, Beijing, China

Gianluigi Ferrari, Università di Parma, Parma, Italy

Manuel Ferre, Centre for Automation and Robotics CAR (UPM-CSIC), Universidad Politécnica de Madrid, Madrid, Madrid, Spain

Sandra Hirche, Department of Electrical Engineering and Information Science, Technische Universität München, München, Germany

Faryar Jabbari, Department of Mechanical and Aerospace Engineering, University of California, Irvine, CA, USA

Limin Jia, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China

Janusz Kacprzyk, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

Alaa Khamis, German University in Egypt El Tagamoa El Khames, New Cairo City, Egypt

Torsten Kroeger, Stanford University, Stanford, CA, USA

Qilian Liang, Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX, USA

Ferran Martin, Departament d'Enginyeria Electrònica, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain

Tan Cher Ming, College of Engineering, Nanyang Technological University, Singapore, Singapore

Wolfgang Minker, Institute of Information Technology, University of Ulm, Ulm, Germany

Pradeep Misra, Department of Electrical Engineering, Wright State University, Dayton, OH, USA

Sebastian Möller, Quality and Usability Lab, TU Berlin, Berlin, Germany

Subhas Mukhopadhyay, School of Engineering & Advanced Technology, Massey University,

Palmerston North, Manawatu-Wanganui, New Zealand

Cun-Zheng Ning, Electrical Engineering, Arizona State University, Tempe, AZ, USA

Toyoaki Nishida, Graduate School of Informatics, Kyoto University, Kyoto, Kyoto, Japan

Federica Pascucci, Dipartimento di Ingegneria, Università degli Studi "Roma Tre", Rome, Italy

Yong Qin, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China

Gan Woon Seng, School of Electrical & Electronic Engineering, Nanyang Technological University, Singapore, Singapore

Joachim Speidel, Institute of Telecommunications, Universität Stuttgart, Stuttgart, Baden-Württemberg, Germany

Germano Veiga, Campus da FEUP, INESC Porto, Porto, Portugal

Haitao Wu, Academy of Opto-electronics, Chinese Academy of Sciences, Beijing, China

Junjie James Zhang, Charlotte, NC, USA

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering - quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

China

Jasmine Dou, Associate Editor (jasmine.dou@springer.com)

India

Swati Meherishi, Executive Editor (swati.meherishi@springer.com)

Aninda Bose, Senior Editor (aninda.bose@springer.com)

Japan

Takeyuki Yonezawa, Editorial Director (takeyuki.yonezawa@springer.com)

South Korea

Smith (Ahram) Chae, Editor (smith.chae@springer.com)

Southeast Asia

Ramesh Nath Premnath, Editor (ramesh.premnath@springer.com)

USA, Canada:

Michael Luby, Senior Editor (michael.luby@springer.com)

All other Countries:

Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)

Christoph Baumann, Executive Editor (christoph.baumann@springer.com)

**** Indexing: The books of this series are submitted to ISI Proceedings, EI-Compindex, SCOPUS, MetaPress, Web of Science and Springerlink ****

More information about this series at <http://www.springer.com/series/7818>

Mykhailo Ilchenko · Leonid Uryvsky ·
Larysa Globa
Editors

Advances in Information and Communication Technologies

Processing and Control in Information
and Communication Systems

 Springer

Editors

Mykhailo Ilchenko
Igor Sikorsky Kyiv Polytechnic Institute
National Technical University of Ukraine
Kyiv, Ukraine

Leonid Uryvsky
Igor Sikorsky Kyiv Polytechnic Institute
National Technical University of Ukraine
Kyiv, Ukraine

Larysa Globa
Igor Sikorsky Kyiv Polytechnic Institute
National Technical University of Ukraine
Kyiv, Ukraine

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-3-030-16769-1 ISBN 978-3-030-16770-7 (eBook)
<https://doi.org/10.1007/978-3-030-16770-7>

Library of Congress Control Number: 2019935980

© Springer Nature Switzerland AG 2019

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This volume is a collection of the most important research results in the fields of information, telecommunication and radioelectronics technologies provided by a different group of researchers from Ukraine in collaboration with scientists from different countries. The authors of the chapters from this collection present in-depth extended research results in their scientific fields.

The volume consists of three parts.

Part I Information Technologies and Research deals with various aspects and approaches to the formulation, analysis and solution of practically important and challenging issues of information systems in general and Internet of things, management services quality control, improved estimates for the reliability indicators, cryptographic technology blockchain, representation of GEO spatial information in particular.

Part II Telecommunication Technologies and Research contains original works dealing with many aspects of construction using research and forecasting of technological characteristics of telecommunication systems.

The presented studies of this part cover a wide range of telecommunication technologies, including over-the-horizon communication systems, satellite communications, multiservice transmission systems as well as in-depth aspects of the study of channel state control algorithms, the implementation of time, and synchronization management technologies in telecommunications.

Part III RadioElectronics Technologies and Research contains relevant papers which show some effective technological solutions that can be used for the implementation of novel systems.

For the convenience of the readers, we briefly summarize the contents of the chapters accepted.

Part I Information Technologies and Research.

The first chapter presented by **A. Baranov** (*Internet of Things: Future Telecommunication*) introduces the possible changes in social patterns of the provision of telecommunication services under influence of using Internet of things (IoT) technology. In particular, some social models of telecommunication operators

and models of the requirements of subjects to receive telecommunication services are presented. The impact of the use of IoT technologies on the activities of telecommunication operators is analyzed. The directions of the possible transformation of the business model of a telecommunication operator that will help meet the needs of users of IoT technology are considered. The features of the construction of the telecommunication infrastructure of the operator in the interests of meeting user needs are considered. It is shown that the use of IoT technologies significantly increases the load on the radio frequency resource. The technical features and legal models of increasing the efficiency of using radio frequency resource in Ukraine are presented.

Chapter *Cryptographic Technology Blockchain and its Applications* by **A. Luntovskyy and D. Guetter** focuses on the problem of using blockchain technologies and their networked cryptographic applications. An important trend since approximately the year 2000 is the use of modern cryptotechnology “blockchain” for acceleration, transparency, and decentralization of financial transactions and as promising digital payment instruments and cryptocurrencies (e.g., Bitcoin, Monero, and Ethereum).

The influencing factors and sources of blockchain cryptotechnology were discussed, the comparison of centralized bank systems vs. decentralized systems was carried out, the mining process for cryptographic currencies, the concept of a public ledger, the validation principles PoW and PoS are represented, as well as the profitability of cryptographic currencies was analyzed. Furthermore, important applications of blockchain cryptotechnology were shown (such as Smart Contracting, Bletchley) as well as the accompanying risks, and their advantages and disadvantages are discussed. In addition, the malicious applications are discussed such as the ransomware (extortion Trojans). Finally, the potential and future perspectives of blockchain cryptotechnology for real business applications are assessed.

Chapter *Model of transdisciplinary representation of GEOspatial information* by **O. Stryzhak, V. Prychodniuk, and V. Podlipaiev** covers intellectual systems for knowledge extraction and their applications for GIS. On the basis of analysis of the text primary structure, which is formed during its lexical analysis, the possibility is determined and the method of structuring natural language texts with the help of the recursive text reduction procedure using the rules presented in the form of lambda expressions is proposed. It is shown the possibility of dynamically forming such rules by users without special training on the basis of the original structure of the texts is already analyzed. An ontological model of an interactive document is proposed, which is intended to analyze the results of text structuring by the expert. The model of ontological GIS application is an interactive document that allows using affine space to display geospatial information upon request. The models of the interactive document and the ontological GIS application provide a high representative level of information available in text documents (in particular, geospatial). The method described in the chapter allows us to form a structured representation of texts to provide the possibility the information processing

contained in the text by automatically or used computer-aided tools. The model of transdisciplinary information representation implementation as an interactive document function allows quick access to large arrays of thematic information, and in combination with ontological GIS application it solves the problem of transdisciplinary representation of geospatial information.

Chapter *Quality Control for Mobile Communication Management Services in Hybrid Environment* by **L. Globa, M. Skulysh, O. Romanov, and M. Nesterenko** investigated the features of the quality control service that servicing of flows in the mobile communication network. The integration of telecommunications systems and a distributed computing environment allows to create a unified hybrid environment for telecommunications services. It has the ability to control the process of information flow services at each stage and ensure compliance with high standards of quality. Providing the quality of service to end-users of communication networks depends on quality control at all stages of the provision of the service. Today, there is a growing need for well-scalable communication systems that meet the needs of end-users due to dynamically changing service structure provided to end-users and constantly changing requirements for service quality indicators with the growth of traffic volumes. Of particular importance are the management systems of service. In this chapter, the features of quality control service, which controls flows of the mobile communication network, are investigated with the use of partial virtualization of network functions. An architectural solution is proposed for organizing service streams in a hybrid environment. It includes a telecommunication network and cloud computing resources that provide servicing of virtualized functions involved in the organization of service streams. The solution for improving the PCRF system and some procedures that allow ensuring quality control of service flows and controlling the computing resources of a hybrid system, whose work affects the quality of servicing of service flows of the system, is proposed.

In *Improved Estimates for the Reliability Indicators of Information and Communication Network Objects with Limited Source Information*, **D. Mogylevych and I. Kononova** present analytical relationships for quantitative assessment of the lower and upper bounds of reliability indicators for telecommunication objects of two classes: a) objects with structural redundancy, in which the connection time of backup elements is so small that it can be almost neglected and b) objects with structural and temporary redundancy, in which connection time is the final (usually random) quantity. The analysis of these relations allowed us to distinguish functionals characterizing the reliability of the objects under consideration, for which exact boundary values were determined. These results made it possible to obtain two-sided assessments of the reliability indicators of the objects of the above classes under conditions when information on the distribution functions of the connection time of backup elements and the recovery time is limited by knowledge of only two initial moments. An important case for practice is also considered, when along with permanent refusals in the equipment of an object, failures may also occur. Analytical relationships have been obtained that allow us to

calculate two-sided assessments of the reliability indicators of objects with structural and temporary redundancy, when equipment failures are taken into account along with refusals. This chapter concludes with a statement of the results that allow us to find two-sided assessments of the reliability indicators of systems with a serial (in the sense of reliability) connection of objects when structural and temporary redundancy is provided.

Part II Telecommunication Technologies and Research.

M. Ilchenko, S. Kravchuk, and M. Kaydenko (*Combined Over-the-Horizon Communication Systems*) present the scientific and technical principles of construction of new hybrid combined over-the-horizon communication systems with the use of reference small-sized stations of tropospheric communication, relaying intellectual aeroplatforms, and artificial formations. These principles are based on the use of new technologies and software-defined and cognitive radio, cooperative relaying, machine-to-machine, effective interaction system, hardware, and application protocol levels.

In *Prospects for the Development of Geostationary Satellite Communications Systems in the World*, **T. Narytnyk and S. Kapshtyk** are concerned with some issues related to the current state and prospects of development of satellite communication systems GEO in the context of the market of satellite communication services and using in the space segment technical solutions. The design of geostationary satellite systems using high-throughput satellites is considered. The development features of systems with high-throughput satellites of various operators are shown. Special attention is paid to the use of high-throughput satellite technologies in Intelsat Epic satellites for C and Ku bands. The advantages of Intelsat Epic solutions are shown. Trends in the development of regular GSO satellites are considered. The combining tendencies of high-throughput and regular satellites in one orbital position are shown.

L. Uryvsky, A. Moshynska, S. Osypchuk, and B. Shmigel (*Comparison of methods for determining noise immunity indicators of a multiservice transmission system*) discuss an approach to analyze the factors that affect the wireless communication line reliability and, as a result, the performance of communication system. Features of productivity determination and bandwidth are described through the characteristics of noise immunity. Using several methods for calculating the probability of a symbol and bit error data is compared and analyzed. The factors for increasing the noise immunity of a telecommunication system are reviewed. A new vector phase method is proposed to accurately determine the noise immunity of multiposition signals. The formulas proposed in the vector phase method provide high accuracy and complete coincidence with the results of the simulation using the proposed method.

In *Research of the control algorithms for a state of duplex communication channel in the conditions of multipath*, **L. Uryvsky, S. Osypchuk, and V. Solyanikova**, discussed some issues related to the features of the DVB-T2

standard technology, which are designed to provide high noise immunity to the information transmission in duplex multiservice high-speed channels. The object of the study is the influence of transformations used in DVB-T2 systems on the noise immunity of communication line. The expediency of borrowing and transferring tools that are used in DVB-T2 into duplex communication line is considered in order to increase the noise immunity of the system. The proposals on increasing the noise immunity of the duplex communication line with the use of DVB-T2 standard transformations are presented. In this paper, the parameter calculation of a typical radio relay link is carried out and the adaptive encoding and modulation scheme for a duplex transmission system based on a DVB-T2 MCS set are developed. The channel models with fading are analyzed, and the choice of the channel model and methods for calculating the noise immunity parameters using the DVB-T2 technology are substantiated.

N. Biriukov and N. Triska (*Time and synchronization in telecoms*) considered some aspects related to the main principles and trends in the field of synchronization in telecommunications. The presented approach is based on the respective ITU-T research and standardization activities and authors' own studies and academic subjects in the Institute of Telecommunication Systems of "Igor Sikorsky Kyiv Polytechnic Institute." A significant part of the results were obtained by the authors during their work at the Ukrainian Research Institute of Communications—UNDIZ (Kyiv, Ukraine). The fundamental definitions of network synchronization and transmission modes and actual requirements for frequency synchronization (FS), phase (PS), and time of day (ToD) synchronization and their metrics are analyzed. The relationships and limits of FS and PS characteristics are proposed. They are based on the FS and PS conception indivisibility. Two-way precise time protocol (PTP) is analyzed subject to time delay random component and frequency deviation. The analysis of both well known and new metrics evaluating the synchronization accuracy was performed. The PTP time stamp exchange was studied using the maximum likelihood method and minimal time interval errors. Results and challenges for further study are summarized in the conclusion.

Part III RadioElectronics Technologies and Research.

Chapter *Software defined radio in Communications* by **M. Kaidenko and D. Roskoshnyi** covers software-defined radio (SDR) systems and their usage in telecommunications. The presented results are based on relevant studies of authors conducted at the Institute of Telecommunication Systems of "Igor Sikorsky Kyiv Polytechnic Institute." Features of the modern SDR systems were analyzed. The architecture of SDR system that is based on the architecture of communication software was substantiated. The architecture that is oriented on the realization of cognitive radio system was proposed. The creation process of SDR system based on SDR transceiver and technologies of system on chip (SoC) using modern means of software-oriented modeling were described in detail. The variants of interaction between processor cores and programmable logic in SoC during creation of SDR

system were considered. The publication is addressed to scientists, graduate students, and developers of radio communication equipment.

H. Avdieienko and Ye. Yakornov, *Application of Spatial Signal Processing by the Form of the Electromagnetic Wave Phase Front in Wireless Communication Systems*, present a novel approach for the spatial signal processing in wireless communication systems. The main purpose of this chapter is to familiarize the readers with such trend of wireless communication system development and improvement as spatial signal processing by the form of the electromagnetic wave phase front. The issues of spatial signal processing practical application are represented by investigations for improving functional capabilities of the phase direction finders and phase systems for coordinate determination. First of all, it is the possibility of distance determining to radio source that locates in the Fresnel region. The features of spatial signal processing application by the form of the electromagnetic wave phase front for the future development of radio relay communication lines, which ensure reuse of frequency band, are shown. The test bed description of the simplex radio relay communication line is presented. This test bed provides frequency band reuse factor of 2, i.e., transmission and reception of two digital radio signals at the rate of 41 Mbit/s in the same frequency band of 8 MHz due to spatial signal processing usage by the form of the electromagnetic wave phase front.

In ***Bridge equivalent circuits for Microwave Filters and Fano Resonance***, **M. Ilchenko and A. Zhivkov** discuss the bridge equivalent circuits for electrodynamic structures whose amplitude and phase–frequency characteristics are similar to so-called Fano resonance. It is shown that electrodynamic structures based on resonators in parallel channels of energy transfer that are not interconnected correspond to the Fano resonance criteria. The processes of degeneration of oscillations and the relationship between partial and nominal oscillations are considered. We analyzed the most important properties of bridge filters—infinite attenuation, extremely long group delay time. Particular attention is paid to the fact that considered properties of bridge circuits are also inherent in the unit cells of electrodynamic metamaterials, such as split-ring resonators.

S. Ivanov (*Fiber-optic and laser sensors-goniometers*) is concerned with some aspects related to the goniometers as sensors for measurement systems. Fiber-optic and laser sensors are used to determine the rotation angle of an object. Those sensors can be used in many areas of science and technology, including telecommunications (measurement of the rotation angle of an antenna). Creation of this class of devices became possible only with the development and improvement of element basis of quantum electronics. The principle of operation of optical goniometers is based on the vortical Sagnac effect. The interest of foreign firms in optical goniometers is based on their potential applications in many fields of science and technology. These devices can replace difficult expensive electromechanical decision in some cases. The Institute of Telecommunication Systems has developed some types of such sensors based on the original engineering solutions.

We would like to sincerely thank the authors of this collection, because without their hard work of preparing good chapters this volume would not have been successfully prepared.

And last but not least, we wish to thank Series Editor, Prof. Janusz Kacprzyk, Polish Academy of Sciences; Dr. Thomas Ditzinger, Executive Editor, Interdisciplinary and Applied Sciences & Engineering; Ms. Swetha Divakar, Project Coordinator, Books Production; Mr. Holger Schaepe, Applied Sciences and Engineering, Editorial Assistant; Ms. Mirnalini Paramasivam, Springer Project Coordinator (Books); Mr. Ravi Vengadachalam from Springer Nature, for their dedication and help to implement and finish this large publication project on time maintaining the highest publication standards.

February 2019

The Editors
Janusz Kacprzyk
Series Editor

Contents

Information and Communication Technologies	
Internet of Things: Future Telecommunication	3
Alexander Baranov	
Cryptographic Technology Blockchain and Its Applications	14
Andriy Luntovskyy and Dietbert Guetter	
Model of Transdisciplinary Representation of GEOspatial Information	34
O. Stryzhak, V. Prychodniuk, and V. Podlipaiev	
Quality Control for Mobile Communication Management Services in Hybrid Environment	76
L. Globa, M. Skulysh, O. Romanov, and M. Nesterenko	
Improved Estimates for the Reliability Indicators of Information and Communication Network Objects with Limited Source Information	101
D. Mogylevych and I. Kononova	
Telecommunications	
Combined Over-the-Horizon Communication Systems	121
M. Ilchenko, S. Kravchuk, and M. Kaydenko	
Prospects for the Development of Geostationary Satellite Communications Systems in the World	146
T. Narytnyk and S. Kapshtyk	
Comparison of Methods for Determining Noise Immunity Indicators of a Multiservice Transmission System	167
L. A. Uryvsky, A. V. Moshynska, S. A. Osypchuk, and B. O. Shmihel	

**Research of the Control Algorithms for a State of Duplex
Communication Channel in the Conditions of Multipath** 186
L. Uryvsky, S. Osypchuk, and V. Solyanikova

Time and Synchronization in Telecoms 205
N. L. Biriukov and N. R. Triska

Radio Engineering and Electronics

Software Defined Radio in Communications 227
M. M. Kaidenko and D. V. Roskoshnyi

**Application of Spatial Signal Processing by the Form
of the Electromagnetic Wave Phase Front in Wireless
Communication Systems**. 239
Hlib Avdieienko and Yevhenii Yakornov

Fiber-Optic and Laser Sensors-Goniometers 262
S. Ivanov

**Bridge Equivalent Circuits for Microwave Filters and Fano
Resonance** 278
M. Ilchenko and A. Zhivkov

Author Index. 299

Information and Communication Technologies



Internet of Things: Future Telecommunication

Alexander Baranov^(✉) 

National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”,
Peremoga Avenue 37, Kyiv 03056, Ukraine
baa_l@ukr.net

Abstract. The article examines possible changes in social patterns of the provision of telecommunications services under the influence of using the Internet of Things Technology. In particular, social models of telecommunications operators and models of the requirements of subjects to receive telecommunications services. The impact of the use of IoT technologies on the activities of telecommunications operators is analyzed. The directions of possible transformation of the business model of a telecommunications operator that will help meet the needs of users of the Internet of Things technology are considered. The features of the construction of the telecommunications infrastructure of the operator in the interests of meeting user needs are considered. It is shown that the use of technologies of the Internet of Things significantly increases the load on the radio frequency resource. The technical features and legal models of increasing the efficiency of use of the radio frequency resource of Ukraine are considered.

Keywords: Internet of Things · Telecommunications · Infrastructure · Models · Radio · Frequency

1 Introduction

With the widespread use of Internet of Things (IoT), according to experts, the number of objects connected to the Internet can reach 100 billion by 2025, much of which will generate a large amount of data that will be transmitted using wireless telecommunications [1]. At the same time, it is predicted that data transfer traffic will grow by three orders of magnitude (1000 times), which will lead to a sharp increase in the relevance of solving the problem of effective use of limited radio frequency resources (RFR).

Today, thanks to the introduction of IoT, there is already a fundamental change in user requirements and expectations for the scope, list and quality of electronic communications services based on radio technologies. This, in turn, requires the transformation of the business model of the operator of electronic communications (operators), which in fact determines the relevance of the study of social models of operators' activities, including the use of RFR.

For several decades, the International Telecommunication Union (ITU), the European Union and the National Telecommunications Regulatory Authorities have made considerable efforts in finding solutions to the problem of ensuring harmonized efficient use of individual RFR ranges. As a rule, earlier these efforts were limited to

carrying out regulatory (legal), organizational, engineering and technical measures. Of course, first of all, the search for improving the efficiency of using RFR is based in the technical direction. But using only technical potential does not allow solving the problem of increasing the efficiency of using RFR fully.

2 Social Models in Telecommunications

Legal regulation problems that have arisen as a result of changes in user requirements and expectations for the scope, list and quality of the electronic communications service (ECS) should methodologically be investigated, like any other legal problems, studying social models and legal models that are a superstructure for these social models. As this takes place, we will assume that the identification and localization of problems, including legal ones, the search for the causes of their occurrence and solutions is more effective if you do research not only within the social system of a certain subject area (telecommunications sphere), but within the social system of higher order hierarchy.

The social model, in our case, that needs to be studied is the social model of the ECS provision system, and the super-system is the social system of the fields of entities activity receiving ECS. Figure 1 shows the social model of the super-system and its component - the social model of the ECS provision system (system 1).

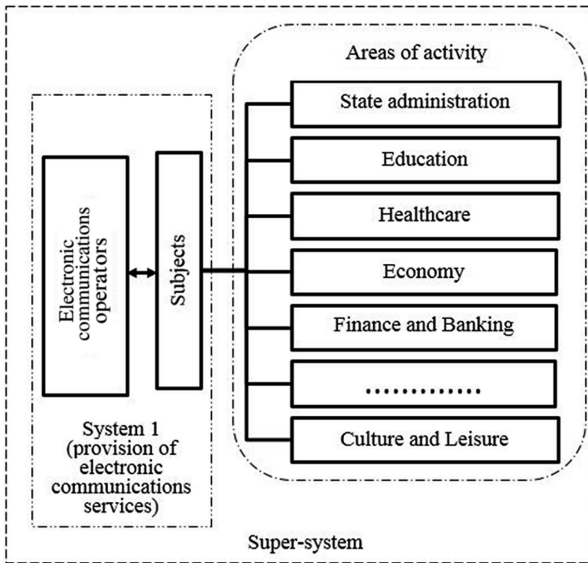


Fig. 1. The social model of the super-system and its component – the social model of the ECS provision system.

In the social model (Fig. 1): the subjects are any individuals or legal entities who are involved in various scope of action and can be consumers of ECS; operators are business entities that, in accordance with national legislation, have acquired the right to provide public ECS, including mobile ones; Scope of action are areas of state, business, professional, personal, and any other activity of entities that can or should be carried out using ECS.

The social model of operators can be presented as a description of their business model as subjects of the electronic communications market using RFR. As this takes place, the content of social processes is mainly determined by the goals and interests both the overall electronic communications market and individual operators or groups of operators, provisions of national legislation, and, in some cases, provisions of international law.

The social model of subjects as ECS users is practically reduced to describing user requirements (expectations) in the nomenclature, content and quality of these services, which are determined by the characteristics of their activities in various spheres of social life.

Over the past 40 years, the activities of operators using RFR, within the framework of the social model (Fig. 1), were mainly focused on provision services to entities using separate end devices (terminals).

By the end of the 90s of the last century, the global market for public ECS, including mobile operators (MO), developed mainly as a market for the provision of voice telephony services. The main distinctive feature of the social model of System 1 (Fig. 1) in these years was the focusing of each MO on its own independence (autonomy) in the provision of services to its subscribers. Taking into account the latter, the business model MO of those times will conventionally be called “autonomous”.

Based on the public availability of ECS, System 1 is a queuing system, namely, a system with the same nomenclature, content and quality of service indicators for all users or their individual large groups. As this takes place, the peculiarities of the activity’s content of subjects in various scope of action do not affect the requirements for the indicators of ECS provision.

Under the influence of the widespread use of computer and Internet technologies, the ITU mobile communication standard 3G (the 3rd generation of mobile communication standard) started to be introduced since 2000, which allowed providing users with a set of services that combines both high-speed mobile Internet access and radio communication technology, which forms a data transmission channel.

The reliability and stability indicators of the network are not critical for the “autonomous” business model overwhelmingly, also as are not critical cases of temporary termination of services for a specific user or even groups of users. Also is not a critical problem of reliable provision of geographically continuous coverage throughout the country for the operator to ensure the possibility of provision services.

Thus, formal description of the “autonomous” business model that dominated the world in the ECS market by the end of the 90s is characterized by the following.

A. In part of using RFR:

- (1) licensing - having a license issued by the regulator, which is the only basis for using the allocated RFR range;

- (2) two licenses - one service: a license for the right to use RFR is issued only to an operator who has a corresponding license for the right to provide services;
- (3) exclusivity - only the operator who is licensed for a specific RFR range can use it;
- (4) full coverage - the requirement to ensure the provision of services throughout the territory for which a license to use RFR is granted;
- (5) national coverage - the aspiration of each operator to have a license for the right to use RFR throughout the country;
- (6) non-rationality - eclectic performance indicators used by the RFR operator.

B. In part of the network of MO:

- (1) exclusivity - the specific network of MO, as a complex of technical means of electronic communications and facilities, is operated by only one operator;
- (2) national coverage - a MO's network for the provision of services by each operator is built throughout the country;
- (3) undemanding - low requirements for stability, reliability and sustainability of the network functioning

C. In part of provided services:

- (1) uniformity - the same quality requirements for services provided by any operator to any user;
- (2) invariance - requirements for service levels quality do not depend on the features of the scope of action of the subjects (service users);
- (3) undemanding - low requirements for continuity, reliability and sustainability of service provision;
- (4) sameness – potentially the same range of services that are provided by all operators for any users;
- (5) equality - relatively the same cost of providing a particular service by different operators, all other things being equal (network capacities, coverage areas, radio technologies, etc.).

At the same time, the subjects' new requirements for the services provided, especially in terms of mobility, led to the need of solving the problem of the possible use of network resources by another operator.

The first example of such "borrowing" of resources can be the use of the capabilities of MO by other operators in the interests of roaming provision, which appeared after 1991 with the introduction of the new mobile communication standard 2G (second generation). As this takes place, the business model of operators was still "autonomous".

Since the beginning of the liberalization process in the field of telecommunications (the nineties), legal opportunities have opened for any operator on a transparent, reasonable price and non-discriminatory basis to the infrastructure of another operator. This situation has stimulated the emergence of a new operator's business model - the Mobile Virtual Network Operator (MVNO) model. The ITU documents give the following definition: MVNO is an operator that provides mobile services to end users, but does not have a license to use its own radio frequency [2].

Thus, the essence of the MVNO business model is to use the network resources of another operator: network or frequency, or both, to provide services.

The development of the MVNO business model was influenced by two trends: market and regulatory. The market trend is explained on the one hand by the operator's desire to get at least some income from their unloaded network or frequency resources by providing them for use (rent) to another operator, and on the other, by the desire to satisfy such requirements of entities that cannot be satisfied within the provision mass publicly available ECS or, in other words, this is about the adaptation of ECS quality indicators for a specific functionally and/or geographically limited segment of the activity of a services consumer.

The regulatory trend is explained by the desire of any National Regulator to increase the efficiency of using RFR by searching for new business models of operators, by forming for the dominant operator's possible additional obligations on connecting third-party MVNOs to their own networks.

Since 2000, the MVNO business model has been applied in different countries with different intensity. Nowadays 4 basic MVNO models are known: Branded Reseller MVNO, Light MVNO, Full MVNO, MVNE (Mobile Virtual Network Enabler). In addition, thanks to new computer technologies, in particular, cloud computing technologies, an innovative fifth model of virtual MVNO (vMVNO) has recently appeared. Components of the network operator's infrastructure in which are implemented on standard computer-network equipment (data centres, servers, storage systems, local and distributed networks, data networks, etc.) [3]. Each of these business models has its own specifics, but all of them are united by the fact that they are focused on the use of MVNO operators network and frequency resources of another operator [3].

Thus, from the beginning of the 2000s, the development of a new approach of using of RFR began - sharing of the RFR part by several operators. We call such a business model of operators and virtual mobile operators (MVNO) "autonomous-common", which means maintaining the basic characteristics of the business model "autonomous" with the creation of legal conditions for sharing certain network or frequency resources provided for use to one operator by another operator.

For the "autonomous-general" business model compared to the "autonomous" business model, the differences are characteristic only in the following.

A. In part of using RFR:

- (1) commonality - the RFR range is used not only by the operator who has a license, but also by another that does not have one;
- (2) full coverage or local coverage - the requirement to ensure the provision of ECS throughout the territory for which a license for using RFR has been issued and the possibility of provision services by another operator in part of the territory of the first operator;
- (3) relative rationality - an increase in the efficiency indicators of the RFR using by the operator due to the greater intensity of its use.

B. In part of the network of MO:

- (1) commonality - a specific MO, as a complex of technical means of electronic communications and facilities, can also be provided by other operators.

C. In part of provided services:

- (1) variability - different cost price of providing a specific service of an operator having an MO network and the right to use RFR, and for an operator using the resources of the first one.

So, for several decades, user requirements for the nomenclature, content and quality of ECS remained virtually unchanged, with the exception of the requirements for the ability to transmit digital data. Moreover, user requirements for these services were formed mainly under the influence of the content of the operators' offers.

However, everything changes drastically with the introduction of IoT, which in fact should ensure the real activity of individuals and legal entities both in providing a wide variety of services and in carrying out works of different content in almost all spheres of human activity. In terms of using IoT the satisfaction of requirements for reliability and stability of the MO network becomes critical, and the requirements for the quality of service provision, including geographic coverage of such services, become quite strict. The number of possible end equipment (mobile devices) involved in IoT technologies with one consumer can be calculated in the hundreds and tens of thousands. Interruptions in the provision of ECS are almost completely excluded; this can lead to a halt in the subject's main activities, such as movement of autonomous road transport.

The modern social model of ECS users can be represented by describing the main user requirements (expectations) in the nomenclature, content and quality of these services, which will be shaped under the influence of the peculiarities of the activities of these subjects while using IoT.

A. In part of system-wide user requirements.

1. Adaptive (flexible) selection of content and quality indicators of ECS in order to optimize the cost of their obtaining.

Impact factors: migrating over the territory concentration of devices or sensors connected to the Internet; daily non-uniformity both in terms of volume and speed of data traffic; daily uneven content and quality of ECS.

Examples of activities: movement of vehicles; logistic tasks; work requiring solar lighting (agricultural, landscape and other work); technologies that have a standby mode (emergency services, emergency assistance), etc.

2. The need for various radio technologies in a certain limited area to provide various user requirements.

Impact factors: tasks of different content, for example, providing cloudy or foggy calculations, in-system or intersystem information exchange, and the like; different requirements for power saving modes; Different requirements for service quality indicators for various devices connected to the Internet.

Examples of activities: agrarian sphere - data collection from stationary sensors with a small frequency (soil moisture parameters, crop maturity, etc.), Data collection from mobile sensors with high frequency (tractors, animals, transport, etc.); medical sphere - data transfer during remote operations or data transfer of remote monitoring of human health indicators, and the like.

3. Accessibility at any point of a local territory (a territory whose borders are determined by the territorial interests of the subject or group of subjects).

Impact factors: dynamic movements in space of executive and measuring devices; the mobility of the components of the production or other process in the activities of the subjects; dynamic displacements in the object space of the subjects activities.

Examples of activities: logistics activities for the movement of goods ordered in a city, and in sparsely populated areas; remote sensing of the Earth using drones or other aircraft; road traffic monitoring.

B. In part of special requirements.

1. High rate and volumes of data transfer.

Impact factors: a large number of devices connected to the Internet; a large number of devices that simultaneously generate significant amounts of data; use of applications requiring large amounts of data per unit of time.

Examples of activities: transportation using autonomously controlled vehicles; implementation of various activities in urban environments; monitoring the movement of goods in terms of megalopolises.

2. High reliability of ECS provision.

Impact factors: a large number of devices connected to the Internet; a large number of devices simultaneously generating large amounts of data; use of applications that require a large amount of data per unit of time.

Examples of activities: organization of industrial production; transport using autonomously controlled vehicles and ships; control of flying objects (drones, light aircraft, aero taxi); traffic management in big cities; infrastructure facility management (urban economy, energy, smart house etc.).

3. Short data transfer delay.

Impact factors: the use of applications that require a high speed of reaction to events, a large number of devices connected to the Internet; a large number of interconnected MO.

Examples of activities: remote surgery, autonomous machines; traffic management in megalopolises.

Especially, it should be taken into account that all these requirements or some of them should be satisfied not for the entire set of ECS user subjects, but only for their individual groups and, as a rule, in a limited area (fields of agriculture or cattle breeding, plants, trade and entertainment complexes, territories of settlements, cross-border highways, etc.).

An analysis of the modern electronic communications market provides the basis for the following conclusion: solving the problem of meeting the fundamental changes in user requirements for the nomenclature, content and quality of these services will be possible only in case of the transition from the mass service system of “undemanding” users to specially created “local” electronic communication ecosystems for individual subjects or their groups. It is specifically created local ecosystems of electronic

communications (LEEC) that will allow forming a flexible exclusive set of services for each individual entity or each group of entities that use IoT to implement various activities.

It is assumed that LEEC will have to meet the highest requirements for reliability and sustainability of operation, especially in part of avoiding interruptions in the provision of services and ensuring a guaranteed high (required) level of quality. LEEC, in contrast to the publicly available electronic communications networks, can be geographically both local and cross-border, oriented to a single user or a group of users and ensure the functioning of IoT complexes and systems, including hundreds and tens of thousands of end devices, which can be used as sensors, actuators, software and hardware systems and complexes, data centers and cloud computing systems and much more.

It means that the Internet, which becomes available with the help of LEEC, must have the features of the “industrial” Internet insofar as its purpose is to ensure the continuous functioning of IoT technologies as the technologies on which one or another user’s activity is based. Of course, the industrial Internet can also be used in the interests of individuals, for example, to provide “smart house” technologies or for a system of continuous monitoring of the health indicators of patients from high-risk groups.

The industrial Internet is the Internet, which has such main features: guaranteed high (required for specific activities) quality of data transmission services, high reliability of this process, which is ensured both by the reliability of MO networks and by the optimal construction of the network topology with the necessary reservation, adaptability to changes in the intensity of data traffic and peak loads.

With the widespread implementation and use of IoT, operators face the challenge of provision services of the so-called industrial Internet by building LEEC, which makes it necessary to fundamentally change the business model of activity.

User requirements in LEEC conditions can be met with greater efficiency by building hybrid electronic communications networks based on the harmonious combination of fixed and mobile networks using a combination of different radio technologies. Also the 5G technology besides meeting the exclusive ECS user requirements arising from the use of IoT, will also meet specific requirements such as low cost, low power consumption by end devices and the ability to support a large number of devices connected to the same base station.

Thus, the telecommunications industry, which until recently operated on the basis of the paradigm of provision public services exclusively on the business-to-consumer (B2C) model requires changes. The consumer was massive with the same user requirements in this model. Business interests motivate operators to notice and take into account the various requirements of individual user groups: youth, students, local businesses, ethnic groups, groups of subjects united by common interests or oriented towards loyalty to some kind of business and so on.

Under these conditions, the best type of business model for MO operators will be: business to group consumer (B2gC). The B2gC business models make it possible to provide a very fast response to changes in the users’ requirements of the subjects, by building new LEEC. In modern conditions, the B2gC business model is most effectively implemented using MVNO, since it creates opportunities for a specific group of

users to provide exclusively LEEC services with improved quality indicators, additional functions, marketing services, and the like.

3 Improving the Efficiency of the Use of Radio Frequency Resource

However, there are serious problems connected with the lack of RFR, which is necessary to ensure acceptable quality of services in the context of continued exponential growth of traffic, especially in the era of the Internet of Things [4]. The reasons for the ineffectiveness of the current state of using RFR could be different: from historical and technical to banal - corruption, but the result was that inefficiency is becoming one of the serious barriers to the widespread IoT technologies implementation.

One of the ways to increase efficiency is flexible general and flexible collective use of RFR because it allows:

- save time to enter innovations into the market;
- small or new operators use innovations in conditions of competition with large market players.

Collective Use of Spectrum (CUS) allows an unlimited number of independent users and/or devices to get simultaneously access to the spectrum in the same frequency range in a specific geographic area with a well-defined set of requirements. A promising direction that has a significant potential for collective use of RFR and which the European Union pays special attention to is the development of the so-called “White Space”, which is understood as the range of RFR available to radio technologies (services, systems) at a certain point in time at a given geographical area on a no disturbance/unprotected basis in relation to primary services or in relation to other services with a higher priority on a national basis [5]. One of the first parts of the spectrum that is best for researching White Space problems is the range 470–790 MHz.

Thus, based on the concept of general access, the CUS model allows reducing legal regulatory restrictions, but on the other hand, this implies a significant responsibility of spectrum users for efficient use of the spectrum and efficient management of obstacles. The CUS legal model should organically correspond the national legal system in part of information infrastructure.

Spectrum Sharing involves the participation of various users who have the right to use a certain radio frequency band, which actually creates potential conditions for each of them in part of access to additional spectrum resources and reduces the barriers to spectrum access for new users [6]. Licensed Shared Access (LSA) is a regulatory approach aimed at facilitating the implementation of radio systems with a limited number of licensees within an individual licensing regime in the frequency band already assigned or expected to be assigned to one or several users. According to this approach, additional users are entitled to use the spectrum (or part of the spectrum) in accordance with the sharing rules included in their license (their right to use the spectrum), which allows all authorized users to provide a certain quality of service (QoS) [7]. For public use, first of all, it is necessary to consider radio frequency bands

intended for special users such as police, fire, border control, military and others whose activities are vital for society.

The problem of business models of operators' activities becomes a serious barrier to the widespread introduction of spectrum sharing, as they are faced with the need to change the logic of business models towards cooperation because competition and cooperation in terms of using the common spectrum become a reality. However, the traditional "autonomous" business model based on competition is practically not adapted to the concept of cooperation.

In general, it can be concluded that a specific business model of sharing a radio frequency resource based on LSA can be determined only if there is knowledge of the features of a particular frequency range, and features of using this range by primary users. Or in other words, the introduction of a regulatory regime for general use of individual RFR ranges will require quite serious and thorough scientific and practical studies of both the principal possibility of using such a mode and defining features of the business model of operators.

But, in any case, as we can talk about the conclusion of agreements between operators to share RFR with various rights to use, the law should provide for the legal possibility of the regulatory authority to influence this situation by the National Regulator, which should be minimal and should not limit the investment activity of operators.

It is very important in forming legal models for using limited RFR is to prevent the conflict of LSA use together with exclusive models of spectrum use. Therefore, LSA should be an additional solution for access to the spectrum in certain ranges, rather than a replacement for the usual exclusive access. That is why we need very flexible and adapted legal models for implementing the LSA approach to spectrum allocation, as this may ultimately change the rules of competition.

The need to build flexible and adapted legal models for the distribution and use of RFR is also determined by significant changes in the business model of ECS provision.

Soon with the massive advent of IoT technologies, not only traditional operators, who exclusively have the right to use individual RFR bands, but completely unexpected actors on the ECS market, will act as subjects provision ECS. For example, it could be: companies that have industrial infrastructure in industry, agriculture, medicine, etc.; service companies (retailers, utilities, logistics companies, etc.); infrastructure organizations serving roads, long-haul pipelines, highways, and the like; companies that organize education, trainings, leisure, especially with large-scale involvement of virtual and augmented reality technology. All this will be the reason for the emergence and application of hybrid business models, which will allow combining services from different business segments [3], which, in turn, will further exacerbate the problem of increasing the efficiency of using RFR due to collective and shared spectrum, as well as its trade in the secondary market.

4 Conclusions

The introduction and use of IoT technologies as a powerful factor of a sharp increase in the effectiveness of any activity in all spheres of social activity requires changes in the existing paradigm of the field of telecommunications.

It concerns the transformation of the business model of electronic communications operators in order to ensure satisfaction of any user requirements of ECS receiving subjects, the introduction of legal models of flexible joint and collective use of RFR.

The implementation of the corresponding transformation of the business model of operators will allow efficient use of investments in the process of dynamic creation of LEEC, which potentially have favourable conditions for the implementation of the industrial Internet concept that is essential for ensuring the continuous and high-quality functioning of the Internet of things.

References

1. Cisco visual networking index: global mobile data traffic forecast update, 2015–2020. Cisco White Paper (2014). <http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-etworking-index-vni/mobile-white-paper-c11-520862.pdf>
2. Virgin Mobile an example of an MVNO. <http://www.itu.int/itu-news/issue/2001/08/mvno.html#top>
3. Shalaginov, A.: Business Model of Convergent MVNO (2017) <https://shalaginov.com/2017/03/31/%D0%B1%D0%B8%D0%B7%D0%BD%D0%B5%D1%81-%D0%BC%D0%BE%D0%B4%D0%B5%D0%BB%D1%8C-%D0%BA%D0%BE%D0%BD%D0%B2%D0%B5%D1%80%D0%B3%D0%B5%D0%BD%D1%82%D0%BD%D0%BE%D0%B3%D0%BE-mvno/>
4. Update on 5G spectrum in the UK. Ofcom (2017). <https://www.boards.ie/b/thread/2057703920>
5. Report on Collective Use of Spectrum (CUS) and other spectrum sharing approaches. Radio spectrum policy group (2011). [http://itlaw.wikia.com/wiki/Report_on_Collective_Use_of_Spectrum_\(CUS\)_and_Other_Spectrum_Sharing_Approaches](http://itlaw.wikia.com/wiki/Report_on_Collective_Use_of_Spectrum_(CUS)_and_Other_Spectrum_Sharing_Approaches)
6. Communication from the Commission to the European Parliament, the Council, the European Economic and social committee and the Committee of the Regions. Promoting the shared use of radio spectrum resources in the internal market. <https://ec.europa.eu/digital-single-market/sites/digital-agenda/files/com-ssa.pdf>
7. RSPG Opinion on Licensed Shared Access (2013). https://circabc.europa.eu/sd/d/3958ecef-c25e-4e4f-8e3b-469d1db6bc07/RSPG13-538_RSPG-Opinion-on-LSA.pdf



Cryptographic Technology Blockchain and Its Applications

Andriy Luntovskyy¹(✉)  and Dietbert Guetter²(✉)

¹ BA Dresden University of Cooperative Education (Saxon Study Academy),
Hans-Grundig-Str. 25, 01307 Dresden, Germany
Andriy.Luntovskyy@ba-dresden.de

² Dresden University of Technology,
Noethnitzer Str. 46, 01187 Dresden, Germany
Dietbert.Guetter@tu-dresden.de

Abstract. The use of modern cryptotechnology “Blockchain” for acceleration, transparency and decentralization of financial transactions and as a promising digital payment instruments and cryptocurrencies (e.g. Bitcoin, Monero, and Ethereum) has become an important trend since approximately 2000.

This article focuses the problem of using blockchain technologies and their networked cryptographic applications and apps.

The influencing factors and sources of Blockchain cryptotechnology were discussed, the comparison of centralized bank systems vs. decentralized systems was carried out, the mining process for cryptographic currencies, the concept of a public ledger, the validation principles PoW and PoS are represented, and profitability of cryptographic currencies was analyzed. Furthermore, important applications of Blockchain cryptotechnology were shown (such as Smart Contracting, Bletchley) as well as the accompanying risks, their advantages and disadvantages were discussed. In addition, the malicious applications were discussed such as the ransomware (extortion Trojans).

Finally, the potential and future perspectives of Blockchain cryptotechnology for real business applications were assessed.

Keywords: Blockchain · Cryptocurrency · Electronic payments · Decentralized payment system · Mining · Energy efficiency · Bitcoin · Ethereum · Monero · Smart Contracting · Ransomware · Profitability

1 Motivation: Payment Instruments in Past and Future

Blockchain is a cryptographically distributed computer network application supporting a decentralized payment system and decentralized financial online transactions in the peer-to-peer (P2P) concept. However, the economic success of this cryptotechnology will be evident in the next 10 up to 20 years.

Figure 1(a, b) depicts the historical development of the payment instruments from archaic shells and early coins to e-cash and cryptocurrencies, which can accelerate financial transactions and significantly reduce the cash mass. The important milestones of Blockchain technology are as follows [1, 2, 5, 11, 12]:

- 1991 – The basic principles from S. Haber and W. Scott: cryptographically secured chaining of individual blocks
- 2000 – Theory for cryptographic blockchaining of Stefan Konst as well as some implementation solutions
- 2008 – White Paper “Bitcoin: A Peer-to-Peer Electronic Cash System” for the conception of a distributed database system BTC created by so-called “Satoshi Nakamoto” (a pseudo of the known developing group as well as a lot of numerous speculations about the developer name, i.a. Elon Musk was mentioned, the founder, CEO and CTO of the companies like PayPal, SpaceX and Tesla)
- 2009 – Launch of the first publicly distributed worldwide Blockchain.

Additionally, an important question on the edge: who did actually invent and create Bitcoin?

Assumption:

Bitcoin was combined by the names of the prominent companies Samsung, Toshiba, Nakamichi and Motorola, i.e.

Satoshi = Samsung + Toshiba

Nakamoto = Nakamichi + Motorola.

A graphical comparison of the decentralized chaining of the secured blocks with a centralized banking system can be seen in Fig. 2 (a, b).



Debentures

a) Central system: issuing authorities, states, banking



 **bitcoin**

Cryptocurrency

b) Decentralized
Blockchain system

Fig. 1. Payment instruments in past and future

The deployment of Blockchain technology speaks mainly for a decentralized financial system. The advantages of such a solution are obvious:

- Sustainability, general transparency and commitment
- Accelerated economic workflows and digitization processes (so-called IT in the digital age)
- Blockchain cryptotechnology is also well-suited to supporting current cryptocurrencies (such as Bitcoin, Monero, Ethereum).

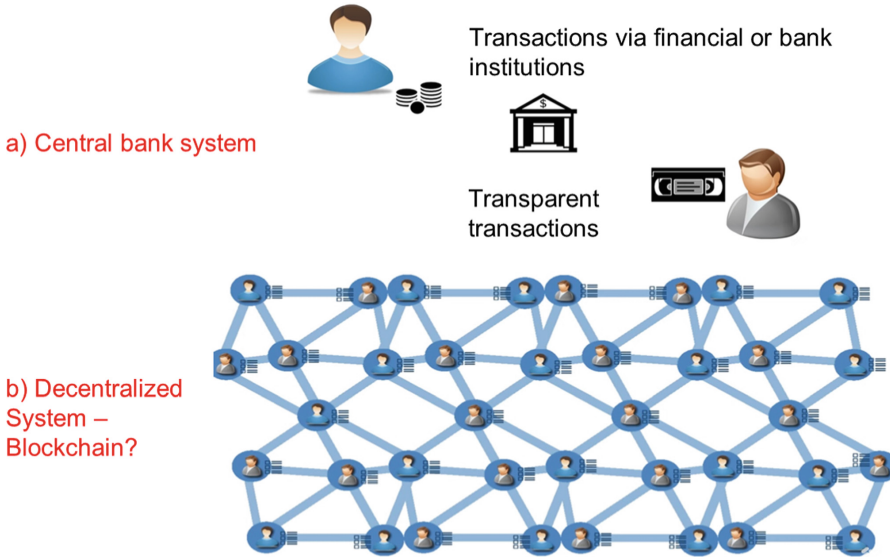


Fig. 2. Decentralized chaining of blocks instead of a central bank system.

2 Blockchain Architecture

Decentralized, cryptographically secured and unified blocks, their chains and transactions are grouped under a general, global public ledger (account), the structure of which is as follows (Fig. 3).

The Blockchain, as a networked Public Ledger, consists of participating nodes that represent an efficient P2P communication model. Typical features of the Blockchain are as follows:

- Redundancy and synchronization
- Cryptographic hash procedures for integrity assurance and attack safety
- Decentralized management and control of the Blockchain
- Network subscribers are also referred to as Nodes (Full-Nodes, Miners, Validators) and run redundantly with mutual synchronization

In addition, large block volumina can cause the “Big Data” problem.

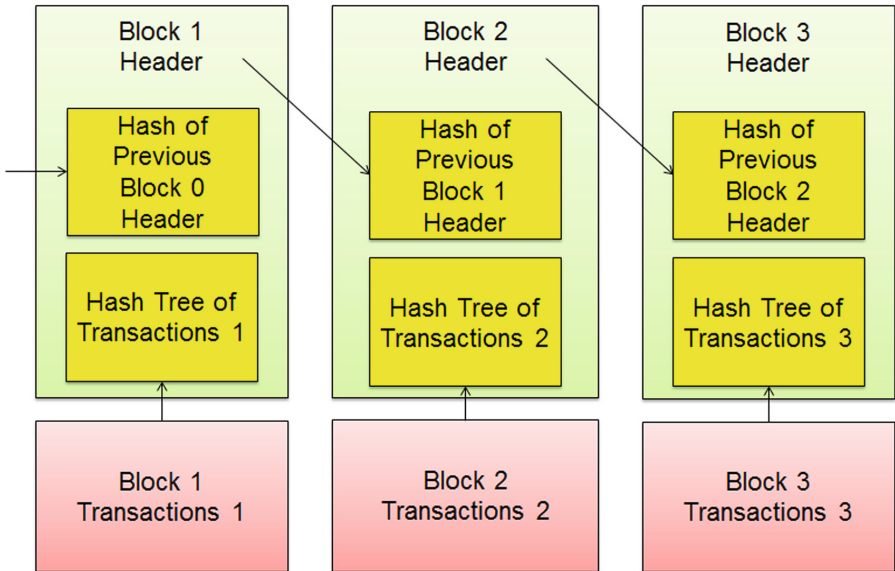


Fig. 3. Block chaining (headers and blocks within a hash tree)

Figure 4 depicts the structure for an exemplary block chain.

- The defining block chain (yellow color) consists of the longest sequence of secured blocks from the origin to the current block (blue).
- Alternative chains (pink color) became orphan as soon as they are shorter than another chain.

Within the Blockchain architecture between the following basic components can be distinguished: the simple Nodes, the Full-Nodes, and Miner/Validator:

1. Nodes:

Each Blockchain participant (computer, smartphones, tablets, or even clusters) is qualified as Node, if he has installed the corresponding software, which runs based on the Bitcoin protocol or the program code of Bitcoin.

2. Full-Nodes:

- A Node with full local copy of the Blockchain
- Checking for so-called “consensus rules”

3. Miner/Validator

- The individual participants or mining pool (high resource requirements regarding hardware and energy consumption)
- Finalising of blocks (Miner – block generation, Validator – proving)
- Externally they act each like a large participant, but in fact many small blocks are generated for payment in fractions of the cryptocurrency units.

BTC Public Ledger (Blockchain)

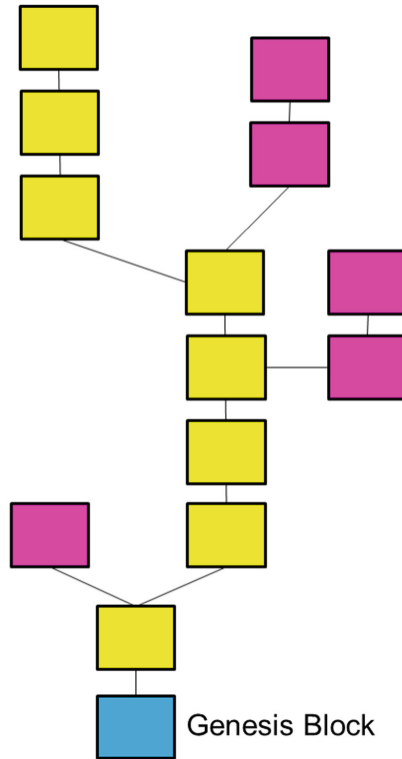


Fig. 4. Distributed digital public ledger

However, the following problems occur during the Blockchain operation:

- Enormous energy consumption due to “mining” of cryptocurrencies (processing of the hash blocks via its algorithmic complexity).
- Exponential memory growth (including capacity migration between USB media, smartphones, PC, storage media such as SAN/NAS, as well as cloud storages)
- Cryptographic data security is guaranteed, but privacy issues may arise. One way out is as follows: no processing the complete Blockchain with all the transactions, but only use of excerpts of the Blockchain without a prehistory.

3 Cryptographic Technology Blockchain and Cryptocurrencies

A cryptocurrency is a digital payment instrument under use of cryptographic methods to realize a decentralized and secure payment system. The following questions can occur for describing of the discussed functionality:

- How is new money created? – Create a new block
- How can the transactions be stored? – Creator of a block selects certain transactions
- How does the respective credit balance come about? – A credit is the sum of all processed transactions of a user.

Example 1

The examples of the cryptocurrencies: Bitcoin, Ethereum, Ripple, Bitcoin Cash, Litecoin. One of them, the cryptocurrency Monero (XMR), has the following properties:

Proof-of-Work with CryptoNight

Memory intensive, but relatively low computational effort

Supported WebAssembly and Coinhive (Coinhive.com) with JavaScript-Mining for Monero-Blockchain

The users can run the miner directly in their browsers and “mine” the XMR for an ad-free experience!

Table 1 contains an overview of current cryptocurrencies.

Table 1. Overview of cryptocurrencies and market growth (sources: [3, 4, 13], <https://coinmarketcap.com/>)

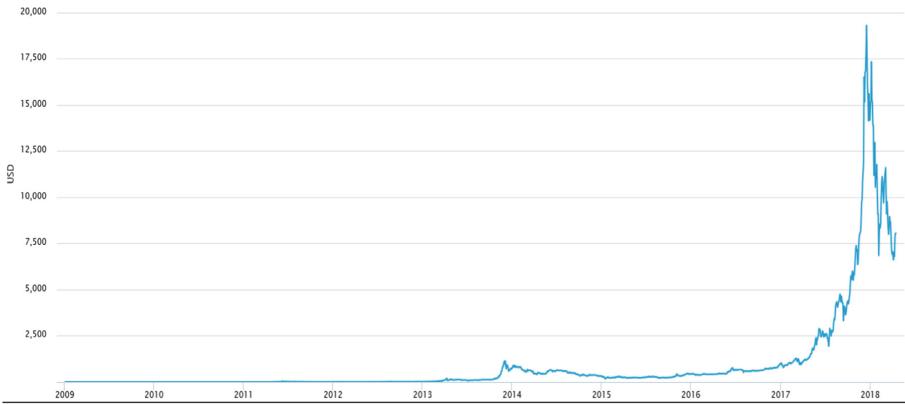
#	Currency	Shorthand symbol	Launch	Mining available	Percentage	Emission
1	Bitcoin	BTC	2009	yes, SHA-256	84,2%	3,9% per anno
2	Ethereum	ETH	2015	yes, Ethash	6,4%	14% per anno
3	Ripple	XRP	2013	no	2,1%	4,1% per anno
4	Litecoin	LTC	2011	yes, Scrypt	1,4%	11% per anno
5	Ethereum classic	ETC	2016	yes, Ethash	0,5%	14% per anno
6	Monero	XMR	2014	yes, CryptoNight	0,5%	20% per anno
7	Dash (formerly darkcoin)	DASH	2014	yes, X11	0,4%	11% per anno

Digital payment instruments under use of cryptographic principles, which built a fast growing market nowadays, are often created via deployment of the hash algorithms titled SHA256 (FIPS NIST 2008). Such algorithms (refer Table 2) were introduced since 2002. The first representant was titled SHA-1. The successor was SHA-2 (including the further modifications like SHA-224, SHA-256, SHA-384, SHA-512, SHA-512/224 and SHA-512/256). The hash size is variable between 224, 256, 384 and 512 bits by the parametrized round number 64 or 80.

Table 2. Overview of NIST hash algorithms [17]

Hash algorithm	Message size (bits)	Block size (bits)	Word size (bits)	Message digest Size (bits)
SHA-1	$<2^{64}$	512	32	160
SHA-224	$<2^{64}$	512	32	224
SHA-256 (blockchain)	$<2^{64}$	512	32	256
SHA-384	$<2^{128}$	1024	64	384
SHA-512	$<2^{128}$	1024	64	512

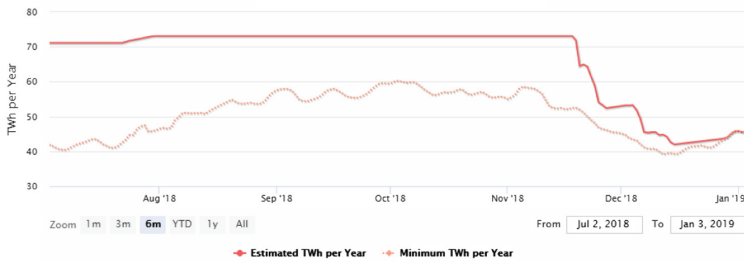
The cryptoapplication Bitcoin (founded in 2009 and based on the Blockchain) is today considered as one of the most important cryptocurrencies (shorthand symbol BTC). The value growth as well as energy consumption by BTC was depicted in the given diagram (refer Fig. 5a, b).



(a) Value growth by BTC

Bitcoin Energy Consumption Index

Bitcoin Energy Consumption Index Chart



(b) Energy consumption by BTC

Legend (b):

- Estimated energy consumption per anno (upper curve),
- Minimum energy consumption per anno (lower curve as a dash-dot line)

Fig. 5. Hype period for the cryptocurrency Bitcoin (Sources: <http://www.heise.de/forum/>, <http://www.statista.de/>, <https://digiconomist.net/bitcoin-energy-consumption>)

However, there is also criticism ([7, 9, 15, 16], <http://www.heise.de/forum/>):

- Blockchain is itself the biggest problem for Bitcoin (due to energy and memory consumption).
- Bitcoin was primarily intended as a decentralized payment system. But due to the exponential growth (storage requirements) of the Blockchain, anyone can still hold the blocks on its local HDD or other storage device. On the smartphones or other mobile devices, the retrieving is also unthinkable anyway. This may affect i.a. the obligation and traceability of transactions.
- Due to critical requirements on ever-growing memory for the Blockchain the users hold the blocks on centralized servers and cloud services and then access it with their online clients. This leads to a certain centralization and is again depending on the central Blockchain servers offered by bank-analog financial institutions and stock markets. The original advantage of decentralization and thus security against manipulation and server failures is no longer available!

4 Profitability of Bitcoin Is Questionable?

As it was shown the cryptocurrency Bitcoin is one of the most popular cryptocurrency examples. Equation (1) describes the profit of cryptocurrency generation. Equation (2) describes the complexity of the mining process via Blockchain [1]:

$$profit = revenue - (cost_{electricity} + cost_{difficulty}) \quad (1)$$

$$cost_{difficulty} = \frac{maximum - difficulty}{current - difficulty} \times \frac{2^{32}}{hashrate} \quad (2)$$

The good profitability of BTC production is difficult to consider or predict due to enormous algorithmic complexity and rising energy prices (Fig. 6).

Example 2

On January 4, 2019 the cryptocurrencies Bitcoin (shorthand symbol BTC) and Monero (shorthand symbol XMR) stood by an exchange price of (refer <https://de.coinmill.com/>):

- Cryptocurrency Bitcoin obtains the value: 3364 EUR = 1 BTC
- Cryptocurrency Monero obtains the value: 43 EUR = 1 XMR.

It should be noted that the variable exchange rates of these cryptocurrencies are not related to the realities of the world industry and trade, so investments in these domains is highly risky. Table 3 contains the most important questions regarding modern cryptocurrencies:

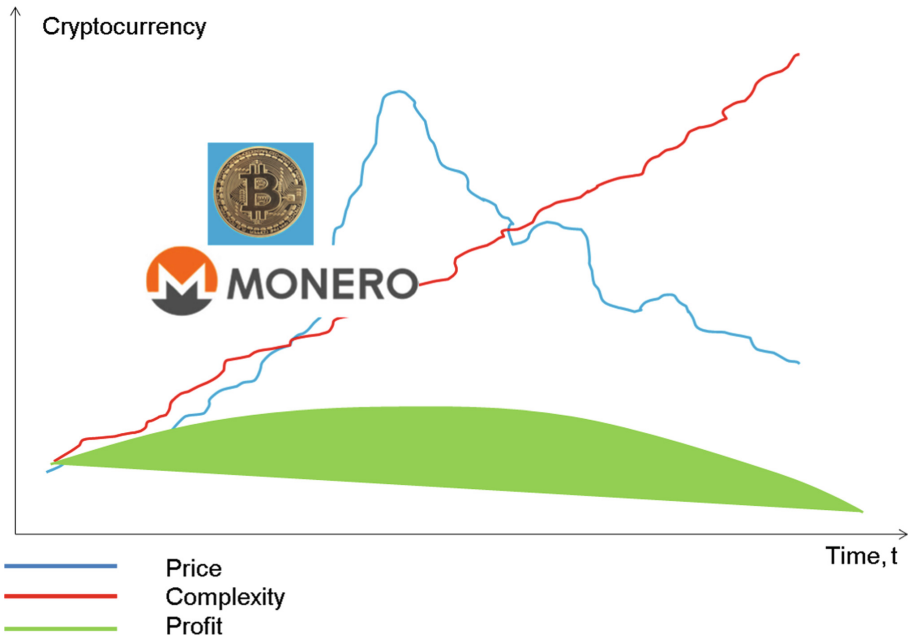


Fig. 6. Profit evaluation for Bitcoin [1, 15]

Table 3. FAQ regarding modern cryptocurrencies

Question	Answer
1. How does it operate?	See above
2. How is new currency generated?	You have to generate a new block for the Blockchain system
3. How can I save transactions?	The blocker chooses certain transactions
4. How is the loan balance provided?	Credit is the sum of all user transactions

5 Operation and Validation of Blockchain

The formation of a transaction block is depicted in Fig. 7. During the Blockchain operation the transaction block is generated accompanied by the both validation processes (Fig. 8):

- PoW, Proof of Work
- PoS, Proof of Stake.

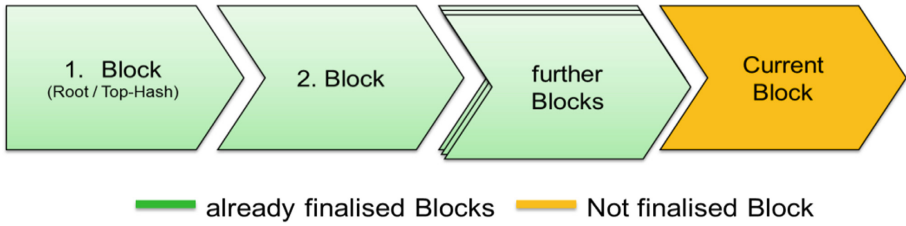


Fig. 7. Blockchain operation: block construction

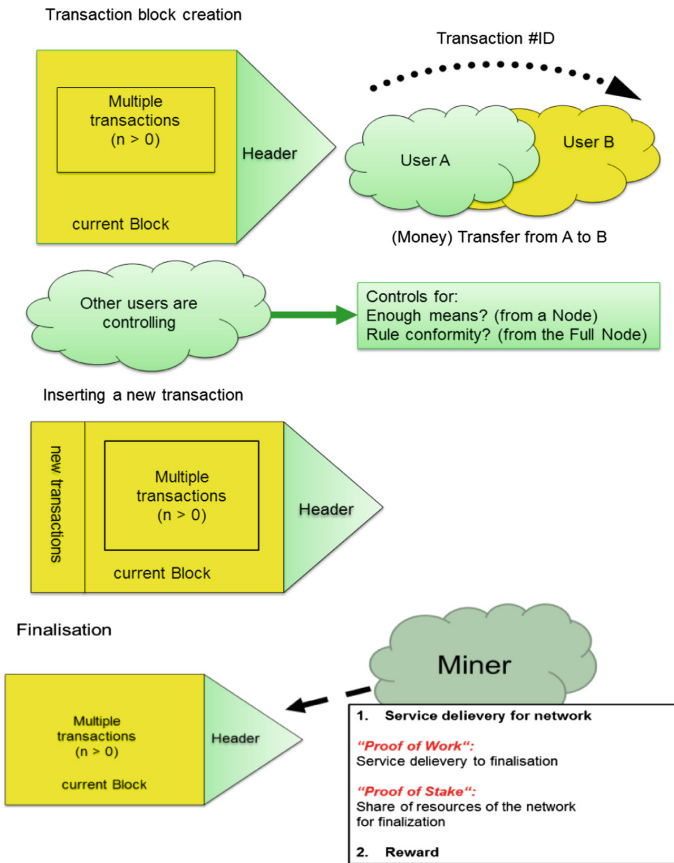


Fig. 8. Blockchain operation: generation of a transaction block

The complete tracking of the chain begins from the first own transaction (refer Fig. 9).

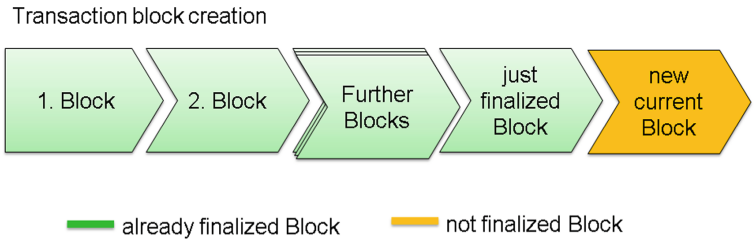


Fig. 9. Blockchain operation: complete tracking of the chain from the first own transaction

What is understood in the frame of the Blockchain technology under the validation processes titled PoW (Proof of Work) and PoS (Proof of Stake) respectively

1. Proof of Work, PoW

PoW is the cryptographic puzzle to confirm the complexity of a particular task and service functions. This demonstrates an overuse of this service, such as a Denial-of-Service-Attack on an Internet host, or a spamming e-mail chain to an SMTP server. PoW should prevent such excessive use of the considered service (Fig. 10).

Proof of Work

„Proof by work” or “proof of employment”
Original idea to avoid the junk emails

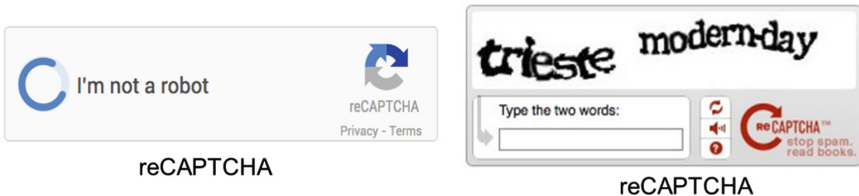


Fig. 10. Blockchain operation: PoW

2. Proof of Stake, PoS

Under PoS is understood a complete or partial proofing method which can create so-called “consensus”: it means which party (which participant) of the Blockchain is allowed to generate the next block (Fig. 11). In this case, a weighted random selection is made, in which weightings are determined for each participant, depending on the duration of the participant’s participation and/or his “assets” (stake). Unlike the PoW, which uses cryptocurrencies Bitcoin and Ethereum, PoS does not consume time or energy like mining processing, nor can the network be “intercepted” because it has a single processing power (“51% attack”) (see Figs. 10 and 11). Instead of a miner, a validator is used whose selection is stochastic.

The best selection options have the validators that stay in the system the longest. A specific part of the Blockchain system is reserved for verification. Such a reserved deposit or “pledge” is distributed and a check is made. In the presence of injuries/offenses, the custody is definitely lost.

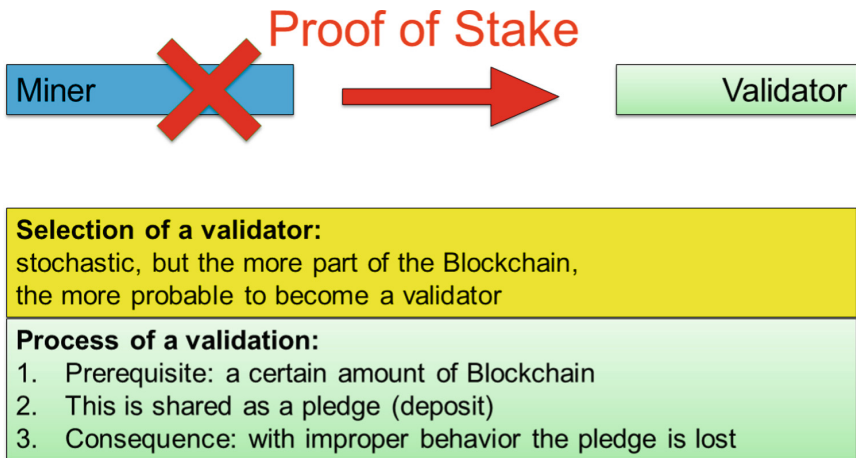


Fig. 11. Blockchain operation: PoS

6 Practical Blockchain Applications

6.1 Framework MS Bletchley

In July 2016 the framework MS Bletchley was launched as a specific service BaaS (Blockchain as a Service) and as a part of the Azure cloud platform. The main goal of the framework introduction [6, 8, 14] was the acceleration of the development of the practical Blockchain applications aimed to financial institutions, manufacturing, retail, healthcare, public sector, media on a common platform. The above mentioned platform contains three following layers (Fig. 12):

1. Base platform tier
2. Middleware (MW) tier
3. Industrial solutions layer.

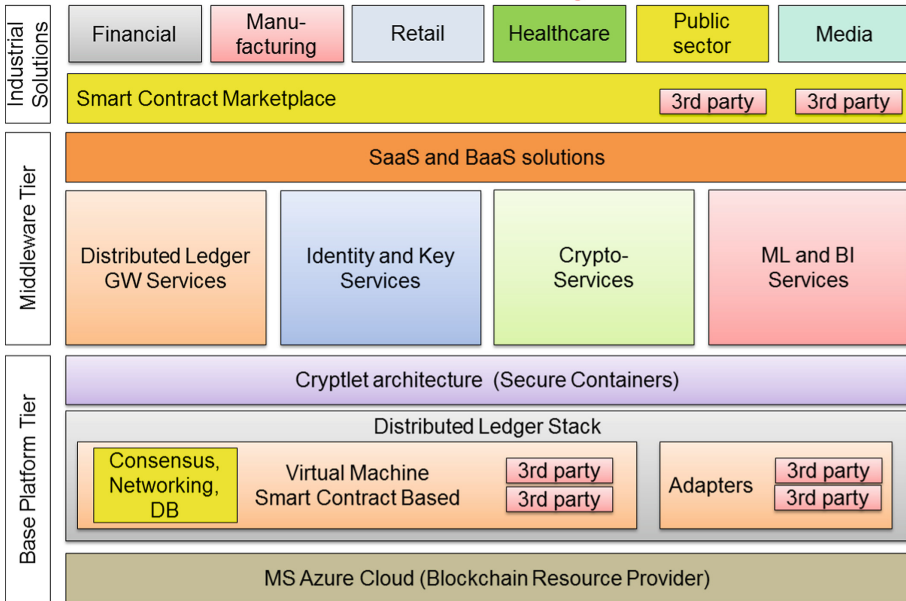


Fig. 12. Blockchain platform MS Bletchley (based on: <http://www.microsoft.com>)

The framework MS Bletchley introduces two new concepts: Blockchain middleware (MW) with so-called cryptlets. To SaaS and BaaS belong i.a. so-called ML and BI functionality (Machine Learning and Business Intelligence). The Blockchain MW components support the core features of services in the clouds, such as identity management, analytics, or machine learning.

Based on the Azure cloud, these components can work with different Blockchain-based technologies aimed to creation of the practical applications. The cryptlets are the building blocks for the Blockchain technology. They are designed to ensure secure communication between the Azure cloud, the MW ecosystem and each specific customer’s technology. The interoperability of the applications with Azure cloud and Azure stack is secured, as well as with the 3rd parties, the clouds like AWS, Google and further private clouds.

6.2 Smart Contracting

Further Blockchain applications are so-called Smart Contracts [7, 20–22]. A “smart” contract is a software-based agreement that allows and can contain a variety of contract terms. In the course of the usual contract processing (transacting), certain linked

actions can be executed automatically if there is a corresponding trigger. The contracts are offered and signed within and via the Blockchain. The evident advantages of the discussed approach are as follows:

- Digitality and legal openness of the platform
- Transparency and time savings.

Example 3

An example of a Smart Contracting application for DB.de [7, 20–22] is given in Fig. 13.

On the example of Deutsche Bahn



Approach:

- introduction of a special DB Coin
- every train has its own wallet and a payment management for itself

The problems:

- a lot of different contracts
- non-optimized processes
- lack of evidence

Fig. 13. Smart Contracting on the example of DB.de [7, 20–22]

Furthermore, the Blockchain is used for Smart Contracting applications for the following world-wide leading companies deployed:

- Walmart, Maersk, Alibaba, CartaSense, Kuehne + Nagel (aimed to logistics, sea freight, stock exchanges, marking of containers)
- Nestlé, Tyson Foods, Unilever (aimed to food delivery)
- Everledger (the registers for diamond certification) and so on.

7 Risks and Hacking

Despite of advances in modern cryptography, viruses and Trojans are however an important threat to the Blockchain technology [18, 19]. Sure, even the Blockchain can also store any data that is often different from financial transactions, transactions and contracts. Such malicious messages can be written to the source transaction addresses

using the scripting language OP_Return. Therefore, the Blockchain can also block the extremistic or pornographic contents.

As an example of such manipulation of the contents of the Blockchain, the Cryptograffiti.info website may be viewed for read and write accesses.

7.1 Ransomware

The next important topic is so-called Ransomware as a dangerous malware type with a blackmail background and the following features:

- Dangerous malware with extortionate action
- Virus-like operation
- Get through prepared attachments to emails
- Exploiting the vulnerabilities of web browsers
- Encryption of passwords and/or files on the hard disk with a powerful 2048-bit-long RSA key
- Prevented or completely blocked encrypted access by the victim to their own data.

Example 4

The ransomware prevents access to the initially encrypted user data. The victim is required to pay a ransom in cryptocurrency units (frequently in bitcoins), about the equivalent of \$ 300–500. The source of the attack or the recipient of the repayment is difficult to track through the use of cryptography. Some examples of this type of Blockchain apps have been listed in Fig. 14.



Fig. 14. Ransomware (screenshots)

One of the well-known examples of ransomware was crypto-trojan WannaCry (approx. 300.000 PCs in 150 countries were affected) as well as Wcrypt, WCRY, WannaCrypt, Wana Decrypt0r 2.0 etc. Many international companies have suffered from these harmful applications.

The GandCrab Ransomware belongs to one of the latest recent risks. The effect of the GandCrab ransomware can be seen in the above figure. An exemplary workflow for the “careless users” was depicted (Fig. 15). Additionally this includes the use of the no-name and well-unstudied development tools (kits). The ransomware also searches for vulnerabilities in Java, IE, Adobe plugins etc.

Unfortunately, the well-known antimalware like ESET, Avira or Kaspersky do not currently have any antidotes.

Example 5

A Windows host receives an e-mail. The subject of the e-mail is titled “Application form to the advertised position – Nadine Bachert/Sarah Summer/Viktoria Hagen/Caroline Smith/Margareth Williams” etc. The application form contains a cover letter, a photo and a CV. The spam filters can classify these as a non-harmful. The malware is packaged as an attachment in a ZIP file. The file contains an executable file. If this file is executed by the user, a symmetric encryption procedure starts. For the key to restore the encrypted host data a ransom is asked. For each user, a unique URL on an anonymizer gate for hidden service is generated. Behind the gate, within the anonymized network the ransom amount and the instructions for payment can be viewed. The amount will be mostly adjusted to the victim.

What is the average case, how do companies solve the problem? The distribution of primary actions for the victims is as follows:

- 78% restore the system simply with a backup
- 20% pay even blackmail money (ransom) to avoid reputation losses
- 2% even try to decrypt themselves.

The advanced firewall systems [1, 23] such as Cisco Umbrella and Check Point R80 contain Anti-Ransomware and can efficiently detect and prevent unauthorized access to the desktop systems, the clouds, the corporate data centers and networks under use of so-called Intrusion Detection/Prevention Systems (IDS/IPS).

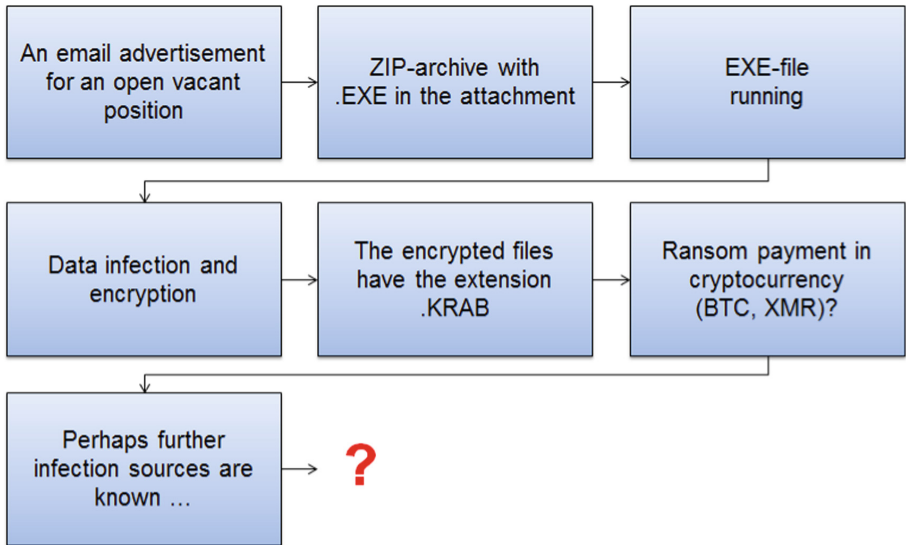


Fig. 15. GandCrab Ransomware: exemplary procedure for “careless users” [19]

7.2 Cryptocurrency Fraud

Some cases of large-scale cryptocurrency fraud in billion amounts became known to the public last time. Cybercriminals use for the robbery the anonymous Internet (so-called Darknet), which can potentially arise through a MIX cascading [1, 23] and is therefore untraceable. This fact effectively ruins the reputation of such MIX networks as Tor, Onion, JAP which are aimed to guarantee primarily the anonymity and privacy for the users. The anonymous Internet is accessible through special software and cannot be generally fully monitored by security analysts.

The attackers can remain anonymous for longer periods. However, the average price for mentioned malware is only \$ 224, with the market value for blockchain-oriented malware now as high as \$ 6.7 million with potential for growth. The attacks on the growing crypto-exchange market now account for up to 25% of all Internet server attacks (according to Computer Forensic experts).

Example 6

According to the statistics for 2018, the following cyber currencies are recognized as the most common targets of fraudsters [18, 19]:

- Monero, XMR (2014) – 44% of hacker quota
- Ethereum, ETH (2015) – 11% of hacker quota
- Bitcoin, BTC (2009) – 10% of hacker quota.

The theft of cryptographic means of payment can be carried out by organized crime cartels or smaller criminal groups. However, it is often about highly skilled professionals looking for an additional criminal source of revenue on the anonymous internet. Most of these incidents related to the theft of the cryptocurrency have been unfortunately recorded in USA (24), China (10) and UK (8).

8 Analysis: What's Stopping the Development of Blockchain?

Despite optimistic forecasts and Blockchain perspectives, development and deployment of the technology is not so fast as the optimists would like to. Certain restrictions prevent this, for example [7, 15, 18]:

1. Control and management

Although the banks and financial institutions recognize the revolutionary character of the Blockchain applications, they also recognize the primary con: when everything is registered on the financial transactions, there is no more banking security and confidentiality! In the case of decentralized transaction processing, the lack of supervisory authorities means i.a. that even weapons and drugs can also be traded using the blocks and cryptocurrencies.

As conclusion:

In some countries – Romania, Kyrgyzstan – a simple way was taken: the Blockchain transactions are prohibited.

The first steps towards the Blockchain regulation were undertaken within the EU. So since May 2018 the General Data Protection Regulation (GDPR, in German: DSGVO) came into force, which restricts the storage and processing of (personal and

privacy) information. The Blockchain has to adhere to these restrictions – at least for EU citizens.

2. Mass character

Despite the apparent increase in cryptocurrency popularity, it should take several years for the cryptocurrency to become fully established. However, insufficient training of the population, skeptical citizens and certain unsuccessful or even criminal inset experiments (e.g. financial pyramid transactions) do not allow the Blockchain to conquer the market immediately.

There is also negative propaganda from some financial institutions and entire states, such as India and China, where the government creates a negative image of the Blockchain and calls every transaction a criminal offense.

3. Costs

The main algorithm for the blockchain is SHA-256. All operations are checked and executed according to this hash algorithm [17]. However, these are very energy intensive processes. By comparison, global hashing exceeded 2018 the level for the power consumption of Cyprus! And these resource requirements will increase even more with a higher number of transactions [3, 4, 13, 20].

9 Conclusions

1. Advantages:

- Blockchain is developing dynamically
- Large IT companies and financial institutions are becoming increasingly involved
- Blockchain is manipulation-proven and theoretically infinitely traceable
- Multiple applications: cryptocurrencies, Smart Contracting and middle-term automation of the domains where an intermediary (confident) is required.

2. Disadvantages:

- Complete replacement of the existing centralized banking system in middle-term is rather hopeless and fantastic; therefore rather niche business as an alternative to centralized banking system and stock exchanges
- Profitability of the Bitcoin and other cryptocurrencies is questionable!
- One thing is evident nowadays:
- The solution to the above listed problems may be too complex and resource-intensive, long-term mass data retrieving (“Big Data”), mass validity tests, prevention of criminal activities (ransomware, fraud)
- So, it is still unclear what way the Blockchain technology will take in the future?

Acknowledgment. Authors’ acknowledgements to the BA Dresden students and scientific fellows (M. Stoll, E. Zumpe, F. Franke, Prof. Dr. L. Zipfel, Prof. Dr. E. Jahn), the colleagues and friends (Prof. Dr. L. Globa from NTUU “Igor Sikorsky KPI” Kyiv, Dr. habil. J. Spillner from ZHAW Zurich) for technical support, inspiration and challenges by fulfilling of this work.




References

1. Luntovskyy, A., Spillner, J.: Architectural Transformations in Network Services and Distributed Systems: Service Vision. Springer, Heidelberg (2017). Case Studies, XXIV, 344p., 238 pict., (ISBN: 9-783-6581-484-09)
2. Staples, M., Chen, S., et al.: Risks and opportunities for systems using blockchain and smart contracts
3. XMR: Was ist Monero? Kryptowährungen und Coins einfach erklärt. <https://www.cryptolist.de/monero/>
4. Monero – Secure, private, untraceable. <https://getmonero.org/>
5. Münzwelt: die Geschichte des Geldes, MDM; Abrufdatum. <https://www.mdm.de/muenzwelt/einfuehrung-ins-muenzensammeln/geschichte-des-geldes/>. Accessed 20 Sept 2017
6. Gray, M.: Projekt Bletchley Whitepaper, MS Worldwide Partner Conference Toronto, July 2016
7. Liesenjohann, M., et al.: Blockchain#Banking: Ein Leitfaden zum Ansatz des Distributed Ledger und Anwendungsszenarien, Bitkom 2016 (Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e. V.), 66 S. <https://www.bitkom.org/noindex/Publikationen/2016/Leitfaden/Blockchain/161104-LF-Blockchain-final.pdf>
8. Rixecker, K.: Modulares Blockchain-Framework von Microsoft: Das steckt hinter Project Bletchley. <http://www.t3n.de/news/checked-c-microsoft-716063/>
9. Malborn Wolff: Blockchain als Open Source: das Hyperledger-Projekt. <https://www.maibornwolff.de/blog/blockchain-als-open-source-das-hyperledger-projekt>
10. Blockexplorer. <https://btc.com/3fb5f4fd6d1129b9797d66be1750b208a10ac2d551125990305211834be94cd9>
11. Bundesanstalt für Finanzdienstleistungsaufsicht: Artikel: Blockchain Technologie; Abrufdatum. https://www.bafin.de/DE/Aufsicht/FinTech/Blockchain/blockchain_node.html. Accessed 1 Oct 2017
12. Welt-Bruttozialprodukt, Bundeszentrale für Politische Bildung. <https://www.bpb.de/nachschlagen/zahlen-und-fakten/globalisierung/52655/welt-bruttoinlandsprodukt>
13. Malmo, C.: Bitcoin hat ein großes Problem: Die Krypto-Währung ist einfach nicht nachhaltig. <https://motherboard.vice.com/de/article/4xa9ed/das-oeko-problem-von-bitcoin-darum-ist-die-krypto-waehrung-nicht-nachhaltig-3920>
14. Oberndorfer, E.: Blockchain as a Service: Microsoft baut einen Blockchain-Marktplatz. <http://www.t3n.de/news/microsoft-blockchain-marktplatz-680639/>
15. Giese, P., Kops, M., Wagenknecht, S., de Boer, D., Preus, M.: Die Bitcoin Bibel – Das Buch zur digitalen Währung. Verlag BTC-Echo (2016)
16. Simon Hülsbömer. Was ist Blockchain? <https://www.computerwoche.de/a/blockchain-was-ist-das>
17. Secure Hash Standards (Algorithms – SHA-1, SHA-224, SHA-256, SHA-384, SHA-512), Federal Information Processing Standards, National Institute of Standards and Technology, 33 p. <http://csrc.nist.gov/publications/PubsFIPS.html#fips180-4>. Accessed October 2008
18. Vishik, C., et al. (ed.): Blockchain: Perspectives on Research, Technology and Policy. <https://pdfs.semanticscholar.org/405a/da981be3a49a78c3e5b281c539938c286001.pdf>
19. Ransomware GandCrab. <https://www.searchsecurity.de/news/252448017/Ransomware-GandCrab-kommt-als-Bewerbung-getarnt>
20. Struckmeier, J.: Cloud and Heat Datacenter for AI, AR and Big Data: secure, scalable and sustainable, Workshop “Energieeffizientes Rechnen und Green IT”, BA Dresden, 59 p, 26 October 2018

21. Karzel, D., Klinger, P.: Prototyping einer Blockchain-Anwendung. In: Stal, M. (ed.) Magazin "Java mit Integrationspektrum" für professionelle Entwicklung und Integration von Enterprise-Systemen, EAN 4194-156107907, pp. 37–40
22. Jansing, M., Tilkov, S.: Kluge Kontrakte auf Basis von Ethereum. In: Stal, M. (ed.) Magazin "Java mit Integrationspektrum" für professionelle Entwicklung und Integration von Enterprise-Systemen, EAN 4194-156107907, pp. 41–45
23. Luntovskyy, A., Klymash, M.: Examination of modern concepts for firewalls and collaborative intrusion detection. *Int. Res. J. Telecommun. Sci.* (2), 44–49 (2016). Issue 14, Kiev (Registration Certificate KB No. 15064-3636P, ISSN 2219-9454, UDC 621.39)



Model of Transdisciplinary Representation of GEOspatial Information

O. Stryzhak^(✉) , V. Prychodniuk^(✉) , and V. Podlipaiev^(✉) 

Institute of Telecommunications and Global Information Space,
Kyiv 02000, Ukraine

sae953@gmail.com, tangens91@gmail.com, pva_hvu@ukr.net

Abstract. The chapter covers Intellectual systems for knowledge extraction and their applications for GIS. On the basis of analysis of the primary structure of the text, which is formed during its lexical analysis, the possibility is determined and the method of structuring natural language texts with the help of the recursive text reduction procedure using the rules presented in the form of lambda expressions is proposed. It is shown the possibility of dynamically forming such rules by users without special training on the basis of the original structure of the texts already analyzed. An ontological model of an interactive document is proposed, which is intended to show the results of text structuring to the expert. The model of ontological GIS – application is an interactive document and allows using affine space to display geospatial information upon request. The models of the interactive document and the ontological GIS - application provide a high level of representativeness of information available in text documents (in particular, geospatial). The method described in the chapter allows forming a structured representation of texts, to provide the possibility of processing the information contained in the text by automatic and automated systems. The implementation of the model of transdisciplinary representation of information as an interactive document function allows quick access to large arrays of thematic information, and in combination with ontological GIS - application, solves the problem of transdisciplinary representation of geospatial information.

Keywords: Lexical analysis · Texture structuring · Ontology

1 Introduction

Providing research of transdisciplinarity is a way of expanding the scientific world outlook, which is to consider a particular phenomenon outside of a single scientific discipline. The development of mechanisms of transdisciplinarity ensures the functioning of the general informational and analytical working environment, and is a priority area for the tasks of informational support of scientific and other research, especially in large integrated projects or in cases where resources are geographically distributed due to the peculiarities of the problems being solved [1–7].

It is advisable to cite five main interpretations of transdisciplinarity (according to Klein) [6]:

- The first approach is connected with the research of the complex unity of the world and the corresponding “difficult” thinking, the expansion of a permanent image of science.
- The second approach defines transdisciplinarity based on the possibilities of a transgressive transition beyond the limits of disciplinary knowledge in sociological and cultural studies.
- The third approach reveals transdisciplinarity as a new form of organization of scientific research, which involves the consolidation of interdisciplinary resources in a single methodological and theoretical framework.
- The fourth approach is the so-called “trans-sector of transdisciplinary”, which focuses on identifying mechanisms for effective interaction between the academic community, the industry and the business sector. In this context, transdisciplinarity can be considered as a resource for synergy between the prospects of expert (academic) and everyday (non-professional) types of knowledge in problem solving.
- The fifth approach is represented by the theoretical concepts of “second type of production of knowledge” and “postnormal science”.

The most general is transdisciplinarity, which is based on the efforts of formal interconnection of the understanding of individual disciplines, which provides the formation of logical meta-frameworks, with the help of which the knowledge outlined in these disciplines can be integrated at a higher level of abstraction than is done in interdisciplinarity. This type of transdisciplinarity is often used in the work of various expert systems and expert groups [8].

Each approach has its own disadvantages and advantages that may arise in solving specific problems. However, the feasibility of transdisciplinary use is undeniable, as evidenced by the text of the “World Declaration on Higher Education for the 21st Century: Approaches and Practical Measures”, adopted by the participants in the International Conference in Paris in October 1998, at the headquarters of UNESCO, in particular its chapter 5 and 6 [7].

In practice, the transdisciplinary approach is used in the form of ontologies [9], which allows for a general scientific classification and systematization of interdisciplinary knowledge.

The most commonly used method for extracting knowledge and meaning from unstructured texts is Deep Learning methods, methods for counting identical words, statistical sequence labeling, supervised machine learning, classification and building ontological models [10–12].

However, the chapter discusses methods related to the construction of ontological models and text structuring together with the use of a dictionary reflecting the meaning obtained in previously structured languages and document models. Particularly relevant are the integration of opportunities associated with extracting meaning from the text and representing this meaning in various geographic information systems.

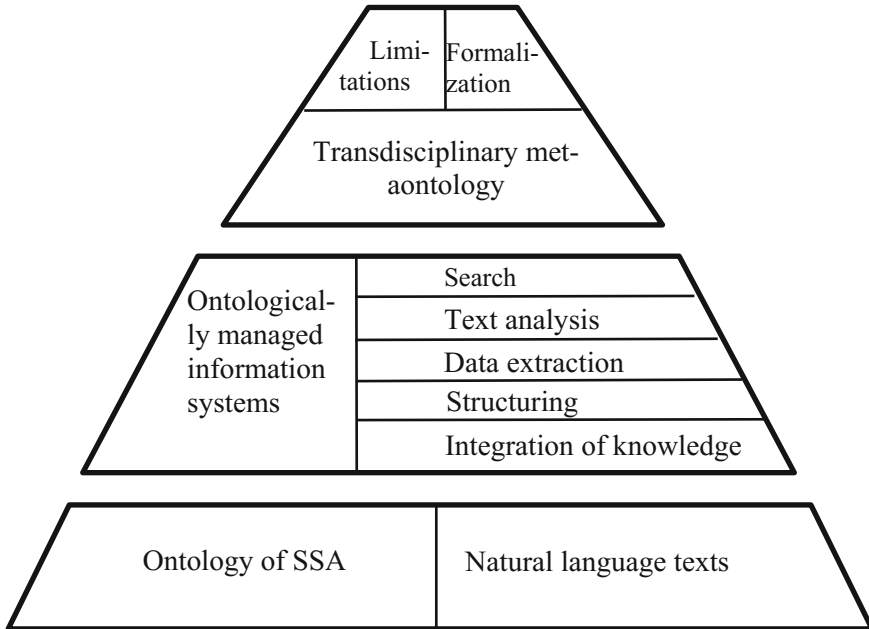


Fig. 1. Interdisciplinary nature of ontologies in scientific research

2 The Concept of an Ontological GIS-Application and Its Formation on the Basis of Natural-Language Documents

GIS is the most natural and convenient way of presenting geospatial information [13–15]. However, the construction of geoinformation system (GIS) can be a rather complicated process if available geospatial data is presented in documents containing weakly structured or even unstructured information. Handling such documents manually can be an extremely labor-intensive process, and the processing of large amounts of such documents is almost impossible [16–18].

Before starting to work with weakly structured or unstructured documents, it is necessary to structure them. During this process, the data is presented in an easy-to-handle form, which can easily be read by standard GIS tools and also conveniently displayed to the end user. This, in particular, may provide an opportunity to find hidden information in the input data [19–26].

The most complicated is the implementation of the structuring of (natural linguistic) NL texts, because this process requires a sufficiently complete formal description of the subset of the language to which they belong. Each of the texts describes a specific subject area (SSA) or a part of it. At the same time, the terms relate to the SSA used in the text form its terminology field. The structuring of the text consists of isolating it from this terminology field, in particular, the identification of the concepts of the corresponding SSA, as well as their attributes and interconnections.

The formed terminology field can be represented using ontology, which is an ordered triple [13, 14, 27–30]:

$$O = \langle X, R, F \rangle \tag{1}$$

where X - the set of concepts of the subject area,
 R - the set of relations between concepts X ,
 F - a set of interpretation functions X and/or R .

Thus structuring of a certain NL text T^T can be represented as a certain transformation (transformation of structuring):

$$F_{str} : T^T \rightarrow O \tag{2}$$

In reality, however, it is not always possible to isolate all the necessary information from the text, therefore, the boundary cases of the formula (1) may occur when the conditions are fulfilled $X = \emptyset, R \neq \emptyset$ or $F = \emptyset$. All possible combinations of these conditions give different variants of ontological constructions, ranging from simple vocabularies to the formal structure of the conceptual knowledge base. In particular, according to [14], one can distinguish:

- (1) $X = \emptyset, R = \emptyset, F = \emptyset$ – unstructured text.
- (2) $X \neq \emptyset, R = \emptyset, F \neq \emptyset$ – glossary;
- (3) $X \neq \emptyset, R \neq \emptyset, F = \emptyset$ – taxonomy;
- (4) $X \neq \emptyset, (R \neq \emptyset | R = R_t \cup R^+), F = \emptyset$ – thesaurus;
- (5) $X \neq \emptyset, R = \emptyset, card(F) = 1$ – simple ontology;
- (6) $X \neq \emptyset, R \neq \emptyset, card(F) > 1$ – an active ontology.

Such a scheme for the classification of ontologies according to the functional attribute is consistent with the description given in [19, 31–47].

Thematic ontology [14] is such ontology in which, besides ensuring the fulfillment of conditions and, the functions of interpretation are added axioms, definition and restriction on the subject of this SSA. The description of all components is presented in some formal language, which can be interpreted by some procedure (algorithm). The scheme of the formal model of thematic ontology is described as follows:

$$O_t = \langle X, R, F, A, D, R_s \rangle \tag{3}$$

where X - the set of concepts (concepts) of the given SSA;
 R - a finite set of semantically meaningful relations between the concepts of SSA;
 F - the finite set of interpretation functions given in concepts and/or relationships;
 A - a finite set of axioms that are used to write always true sayings (definitions and restrictions) in terms of the subject of SSA;
 D - a set of additional definitions of concepts in terms of the subject of the SSA;
 R_s - a set of constraints that determine the scope of conceptual structures of a specific topic of the SSA.

Thematic ontology is a formal representation of conceptual knowledge of the subject area and can be represented by a certain information system. The process of constructing such an information system can be represented as a composition of certain statements, judgments, statements, terms-concepts and relations between them, and its result is the basis for constructing an integral part of the scientific theory - the ontological knowledge base in a given subject area, described in declarative form [21, 22, 48, 49]. Such a scheme for the classification of ontologies by function is consistent with the description: “Ontology or conceptual model of the subject area consists of a hierarchy of concepts of the domain, the links between them and the laws that operate within the framework of this model.”

However, the structure of the SSA, described by the text, is generally limited by the concepts ($X \neq \emptyset$), the links between them ($R \neq \emptyset$). That is, based on texts you can build a glossary O^1 and a taxonomy O^2 . The taxonomy constructed will look like (4).

$$O^2 = \langle X, R \rangle \quad (4)$$

where X - the set of concepts given by the SSA;

R - a finite set of semantically meaningful relations between the concepts of SSA.

However, according to [14], all non-empty products of subsets X and R form a set of actions $F_i \subset F$. Therefore, a plurality of interpretation functions can always be formed when the conditions $X \neq \emptyset$ and $R \neq \emptyset$ are fulfilled and performed for the formula (4). Thus, any formula of the form (4) will always correspond to a certain formula of the form (1).

In fact, structuring text conversion (2) is a multi-stage process, each stage of which requires the use of specialized models and procedures. The general scheme for performing a transformation can be represented by the formula of the form (5).

$$T^T \rightarrow T_{sn} \rightarrow O^1 \rightarrow O^2 \rightarrow O \quad (5)$$

where T^T - NL text;

T_{sn} - the primary structure of the text;

O^1 - glossary;

O^2 - taxonomy (4);

O - ontology (1).

In this case, the first stage, lexical analysis, $T^T \rightarrow T_{sn}$ requires the use of a significant amount of linguistic information, and therefore performed a separate system - a lexical analyzer [32, 40, 50–56].

Obtaining a structured representation of the text in the form of ontology (1) allows it to perform its automatic processing of information in it by means of one or another facilities, but for experts such representation is not appropriate. Because the activity of experts in the corporate information environment can be presented in the form of a system {action -> results}, then it is expedient to submit the information available in

the structured document to the expert with the help of a system of a similar structure. A system of this type can be defined as natural – SN [14, 57].

The natural system (NS), used to display a particular document, provides an interactive interaction with its content. Such document can be defined as interactive and presented with a pair (6).

$$\langle O, SN \rangle \quad (6)$$

where O - ontology, which is a structured representation of a certain document;
 SN - NS, built on the basis O .

The partial case of an interactive document is an *ontological GIS application*, which is an interactive document, NS of which is based on an affine space [58]. The affine space serves as a mathematical basis for the display of geospatial information, and therefore such an NS is able to naturally represent to the expert selected from the text in the process of structuring information of this type. The GIS-application formed on the basis of such an NS can be represented by the formula (7) [13, 58–60].

$$\langle X_g, R_g, A^{op}, S, O \rangle \quad (7)$$

where X_g - the set of geographical entities over which analytical operations are performed to solve the problem;

R_g - a plurality of relationships between geographic objects that determine the type of executed operations;

A^{op} - a set of analytical operations on geographic objects performed in the process of solving a problem;

S - the set of states of the task, which are visualized on the map in the process of its solution;

O - ontological description of geographic objects, processes and tasks of a given subject area.

3 Structuring the Text Using the Method of Recursive Reduction

The process of structuring the text (5) can be divided into two main stages: a lexical analysis $T^T \rightarrow T_{sn}$ that forms the primary structure of the text, and the formation of an ontology $T_{sn} \rightarrow O$, which allows you to select the necessary information from the primary structure and present it as an ontology. For the implementation of the second stage, a method of recursive reduction is proposed, which involves the sequential transformation of the primary structure with the help of a certain set of dynamically-specified user rules.

3.1 Graphic Representation of the Primary Structure of the Text

The primary structure of the text T_{sn} is the result of the work of the lexical analyzer and the input data for the process of recursive reduction. The primary structure contains a structured representation of lexemes (words or symbols), as well as syntactic relationships between them. This structure, in essence, is an oriented graph, and lexemes are the vertices of a given graph.

Any NL text T^T is represented by a plurality of lexemes L , on which the relation of forwarding is determined as \prec . This relation converts L into a linearly ordered set. Also, the text T^T can be represented as a sequence of sentences S , which also defines the relation of the transference:

$$T^T = \{S_1 \prec S_2 \prec \dots \prec S_{n_s}\} \quad (8)$$

where n_s - the total number of sentences in the text.

Each sentence S_i , in turn, is represented by some subset of lexemes:

$$L_{S_i} = \{l_{ij}, j = \overline{1..n_i}\} \quad (9)$$

where n_i - the number of lexemes in the i sentence.

Obviously, the condition is fulfilled:

$$\forall l_1 \in S_1, \forall l_2 \in S_2, S_1 \prec S_2 \Rightarrow l_1 \prec l_2 \quad (10)$$

where S_1, S_2 - arbitrary sentences of the text;

L_1, L_2 - lexemes.

Each lexeme, in its turn, has a number of features:

$$l_{ij} = \langle l_{ij}^T, P_{ij} \rangle \quad (11)$$

where l_{ij}^T - the text representation of lexemes l_{ij} ;

P_{ij} - signs of lexemes l_{ij} .

A lexeme can be linked to other lexemes using syntactic relationships $r_{sn} \in R_{sn}$:

$$r_{sn} = \langle l^1, l^2, k \rangle \quad (12)$$

where l^1, l^2 - lexemes, between which there is a connection;

k - type of connection.

Thus, the oriented graph, representing the primary structure of the NL text, has the form (13)

$$T_{sn} = \langle L, R_{sn} \rangle \quad (13)$$

The main problem of the structure (13), and in particular (11), is the ineffectiveness of working with the text representation of l_{ij}^T lexemes, which is redundant and requires the construction of specialized functions defined on a plurality of text representations of words. Such functions are cumbersome and ineffective, and in software implementation - often depend on the peculiarities of the processing of text variables in a given programming language.

As the set of text representations of lexemes, obviously, is counted, it is possible to construct the transformation of the form (14) [62].

$$V : L^T \rightarrow \mathbb{N} \quad (14)$$

where L^T - the set of text representations of lexemes.

Let the text be written in a certain alphabet W , the number of characters in which $n_W = \text{card}(W)$. This alphabet can be considered as a system of calculus with the basis n_W . Accordingly, each letter $w \in W$ can be put in correspondence with a certain number $i_w \in \mathbb{N}$, which is the index of the given letter in the alphabet. Any word for the input text is a sequence (15).

$$l^T = \{w_1, w_2 \dots w_{n_i}\} \quad (15)$$

where n_i - the length of the word $n_i > 0$;

w_i - the letters of the alphabet W .

If we consider letters w_i as digits of numbers in the corresponding numerical system, then such a number can be converted into a decimal system of the calculation using formula (16).

$$V(l^T) = i_{w_1} \times (n_W)^{n_i} + i_{w_2} \times (n_W)^{n_i-1} + \dots + i_{w_{n_i}} \times (n_W)^0 \quad (16)$$

where i_{w_j} - the index of letters in the alphabet W ;

n_W - the number of characters in the alphabet W .

With the help of the function V , you can replace all l^T the corresponding ones $l^V = V(l^T)$. As a result of this operation, you can get a more effective representation (17) of the set of lexemes.

$$\langle l^V, P \rangle \in L^V \quad (17)$$

where l^V - the code representation of lexeme l ;
 P - grammatical characteristics of lexeme;
 L^V - a set of code representations of lexemes.

In the future, you can consider L^V as a set of lexemes L .

3.2 The Method of Recursive Text Reduction

According to (5), the structuring process of the text performed by the recursive reduction method has the structure (18) and transforms the primary structure of the text (13) into the ontology (1).

$$T_{sn} \rightarrow O^1 \rightarrow O^2 \rightarrow O \quad (18)$$

The method of recursive reduction consists in recursively executing the process of reduction of the input NL text, which consists in applying a specialist operator to it (19).

$$F_{rd} : T_{sn} \rightarrow O \quad (19)$$

The reduction operator (19) is described in terms of the λ -theory.

λ -theory (λ -calculus) is a formal theory developed for the formalization and analysis of the notion of computability [63]. The basic concepts of λ -theory are application (application of a function to an argument) and abstraction. Abstraction means that if $t(x)$ - formula which may contain a free variable x , then the record $\lambda x.t(x)$ is a function f that converts the meaning a to value $t(a)$, in other words, equality (20) holds.

$$(\lambda x.t(x))a = t(a) \quad (20)$$

where x - a certain variable;
 t - formula containing, possibly, the entry of a variable x ;
 a - an argument that defines the value x .

The central element of λ -theory is the notion of the term [63, 64]. The set Λ , λ -terms is determined inductively, as shown in formula (21):

$$\begin{aligned} x &\in \Lambda \\ M \in \Lambda &\Rightarrow (\lambda x M) \in \Lambda \\ M, N \in \Lambda &\Rightarrow (MN) \in \Lambda \end{aligned} \quad (21)$$

where x - arbitrary variable;
 M, N - arbitrary terms.

Besides, in the λ -theory, the so-called Steinfinkle observation (22) operates, which allows us to reduce the functions of many arguments to the function of one argument [63].

$$\lambda x_1 \dots x_n . M = \lambda x_1 (\lambda x_2 (\dots (\lambda x_n . (M)) \dots)) \quad (22)$$

where x_i - arbitrary variables;
 M - arbitrary term.

Important is the notion of β -reduction, which can be defined as the relation (23) between two terms:

$$\beta = \{((\lambda x . M)N, M[x := N]) \mid M, N \in A\} \quad (23)$$

where x - arbitrary variable;
 M, N - arbitrary terms;
 $M[x := N]$ - formula obtained by substitution N instead of a variable x in expression M .

The reduction operator (19) is a combination of four operators (24). while three of the operators perform the conversion steps given by the formula (18), and one performs the auxiliary function.

$$F_{rd} = F_{l*} \circ F_x \circ F_{smr} \circ F_{ct} \quad (24)$$

where F_{l*} - the aggregation operator (25), which performs the auxiliary function, transforming the set of lexemes L into a plurality of constructs L^* . Constructs are a special form of lexemes, and combine the sequence of words or characters, in particular, the phrase. The peculiarity of constructs is that from the point of view of further processing, they can be considered as lexemes (that is, as a single word or a symbol). In this way, the plural $L \cup L^*$ can be used in the same cases as the plural L , in particular, as the input for the operator

$$F_{l*} : L \cup L^* \rightarrow L^* \quad (25)$$

F_x - the operator (26) of the concept X identification.

$$F_x : L \cup L^* \rightarrow L \cup L^* \cup X \quad (26)$$

F_{smr} - operator (27) of the identification of ontological connections R [65], which are divided into the relationships between concepts R_{sem} and auxiliary links between the concept and its contexts R_{sem}^* .

$$F_{smr} : \langle L \cup L^* \cup X, R_{sn} \rangle \rightarrow \langle L \cup L^* \cup X, R_{sn} \cup R_{sem}^* \cup R_{sem} \rangle \quad (27)$$

F_{ct} - operator (28) identifying contexts, which in the future act as attributes of concepts A_x .

$$F_{ct} : \langle X, R_{sem}^* \rangle \rightarrow A_x \quad (28)$$

The total transformation F_{rd} has the form (29), where $R = R_{sem}^* \cup R_{sem}$

$$F_{rd} : \langle L \cup L^* \cup X, A_x, R_{sn} \cup R \rangle \rightarrow \langle L \cup L^* \cup X, A_x, R_{sn} \cup R \rangle \quad (29)$$

In the first step $L^* = \emptyset$; $X = \emptyset$; $A_x = \emptyset$; $L_{sem}^* = \emptyset$; $L_{sem} = \emptyset$, that's why transformation (29) is degenerate into the transform (30). However, with each step, the data sets are replenished with new elements.

$$F_{rd} : \langle L, R_{sn} \rangle \rightarrow \langle L \cup L^* \cup X, A_x, R_{sn} \cup R_{sem}^* \cup R_{sem} \rangle \quad (30)$$

To fully analyze the text, it is necessary to use the operator (24) recursively, for which the operator of the fixed point is used in the λ -theory [63]. To do this, you need to build an auxiliary function F' :

$$F' = \lambda f x. \begin{cases} fF_{rd}x, & F_{rd}x \neq x \\ x, & F_{rd}x = x \end{cases} \quad (31)$$

We apply a fixed point operator Y to it (33). A term YF' is a fixed point for a transformation given by a function F' , that is, for it the relation is satisfied (34). By using the relation (34), the recursive execution of the function (31) can be performed until it reaches the completion condition (35), that is, until the plurality of current data (32) becomes a fixed point for the transformation F_{rd} .

$$\langle L \cup L^* \cup X, A_x, R_{sn} \cup R_{sem}^* \cup R_{sem} \rangle \quad (32)$$

$$Y = \lambda f. (\lambda x. f(xx)) \lambda x. f(xx) \quad (33)$$

$$F'(YF') = YF' \quad (34)$$

$$F_{rd}x = x \quad (35)$$

Condition (35) can be interpreted as “application F_{rd} no longer identifies new information”.

3.3 The Multiplicity of Transformation of Recursive Reduction

Constructs L^* allocated during the transformation (25) are formed using a specialized operation. As a rule, the concept described in the input text is represented in it by a

certain phrase, which in turn is represented by a certain subset of lexemes $\tilde{L} \subset L$. Above this set is a specialized conversion (36).

$$\tilde{L} \rightarrow \langle V^+ (\tilde{l}_1^V, \tilde{l}_2^V \dots \tilde{l}_n^V), P_M(\tilde{L}^*), L^{im} \rangle \quad (36)$$

where n - the number of lexemes in \tilde{L} ;
 l^V - code representation of lexeme l ;
 V^+ - operation combining code representations by lexemes (37);
 $M(L)$ - operation finding the main lexeme of the set L ;
 L^{im} - generating set ($L^{im} = \tilde{L}$).

The operation V^+ is given recursively (37).

$$V^+ (l_1^V, l_2^V \dots l_n^V) = l^V \times [log_{n_w}(V^+ (l_2^V, l_3^V \dots l_n^V)) + 1] + V^+ (l_2^V, l_3^V \dots l_n^V) \quad (37)$$

where $[]$ - operation finding the whole part of the number;
 n_w - the number of characters in the alphabet W , which is used for the text representation of the lexemes.

Conversion V^+ is an analogue of applying transformation V to source text formula of lexemes (38).

$$V^+ (l_1^V, l_2^V \dots l_n^V) = V(V^{-1}(l_1^V) + V^{-1}(l_2^V) + \dots + V^{-1}(l_n^V)) \quad (38)$$

where V - the operation of forming a code representation of lexeme;
 V^{-1} - back to the operation V , $V \circ V^{-1} = V^{-1} \circ V = \lambda x.x$;
 $+$ - combining operation (concatenation) of text representations.

Constructs can also be created on the basis of a combination of constructs and lexemes ($\tilde{L} \subset L \cup L^*$). The main difference is the formation of a generating set L^{im} with the help of formula (39).

$$L^{im} = \{l | l \in L \cap \tilde{L}\} \cup \bigcup_{l^* \in \tilde{L}} L_l^{im} \quad (39)$$

where L - the set of lexemes of the input text;
 \tilde{L} - set of lexemes and constructs from which the construct was formed;
 l - lexemes belonging to \tilde{L} ;
 l^* - Constructs belonging to \tilde{L} ;
 L_l^{im} - forming a set of constructs l^* .

The concept of a generating set L can also be extended to a plural by taking (40).

$$l \in L \Rightarrow L_l^{im} = \emptyset \quad (40)$$

On the set of lexemes L , the relation of a strict linear order \prec is determined. This relation can be extended to a plurality of L^* :

$$\forall l_1 \in L_{l_1}^{im}, \forall l_2 \in L_{l_2}^{im}, l_1 \prec l_2 \Rightarrow l_1^* \prec l_2^* \quad (41)$$

where l_1, l_2 - lexemes;

$L_{l_1}^{im}, L_{l_2}^{im}$ - forming sets of constructs l_1^*, l_2^* .

But in the plural L^* relation \prec , in the general case, will not be a linear order relation, because the condition (42) may be fulfilled.

$$\exists l_1 \in L_{l_1}^*, \exists l_2, l_3 \in L_{l_2}^*, l_2 \prec l_1 \prec l_3 \quad (42)$$

where l_1, l_2, l_3 - lexemes;

$L_{l_1}^{im}, L_{l_2}^{im}$ - forming sets of constructs l_1^*, l_2^* .

The ratio (41) can be trivially extended to $L \cup L^*$. Also, the set $L \cup L^*$ has a relation R :

$$l_1 \in L \cup L^*, l_2 \in L^*, l_1 \in L_{l_2}^{im} \Rightarrow l_1 R l_2 \quad (43)$$

The ratio R , as well as the relation of follow-up \prec , are relations of a strict partial order on $L \cup L^*$.

Now consider the more detailed combination of transformations (25) and (26), in the first step of reduction, when it has the form (44).

$$F_{l^*} \circ F_x : L \rightarrow L \cup L^* \cup X \quad (44)$$

This transform converts each lexeme $l \in L$ into a lexeme itself, as well as a certain set of constructs or concepts that are formed on the basis of this lexeme. If we take as a starting set the set of constructs L^* , then this correspondence can be represented as a hyperratio $G_{im} : L^* \rightarrow L$:

$$l \in L, Y = \{l^* | l \in L_{l^*}^{im}\} \Rightarrow Y G_{im} l \quad (45)$$

where l - lexeme;

l^* - arbitrary construct;

$L_{l^*}^{im}$ - forming a set of construct l^* .

This hyperratio can be extended to a set of concepts X by condition (46).

$$l \in L, Y = \{x | \exists l^*, x(l^*) = x \wedge l \in L_{l^*}^{im}\} \Rightarrow YG_{im}l \quad (46)$$

where l - lexeme;

x - an arbitrary concept;

l^* - arbitrary construct;

$x(l^*)$ - operation of forming a concept based on a construct l^* ;

$L_{l^*}^{im}$ - forming a set of constructs l^* .

Using the formulas (40), (45) and (46), this relation can be trivially extended to $G_{im} : L \cup L^* \cup X \rightarrow L \cup L^* \cup X$.

Hyperratio G_{im} shows the plurality of the ratio of recursive reduction and can be used as an additional mechanism for the formation of ontological links R between the concepts of SSA.

3.4 The Structure of the Reduction Operator and Its Components

As already mentioned above, operator (24) is a composition of operators. Each of them performs one conversion step (5). The full cycle of this transformation structures a portion of the information contained in the input text, after which the conversion is recursively called again until all information is selected. However, each of the components of the reduction operator can also be divided into constituents.

In the general case, the operator of the transformation F is given by the base of the rules G_R for the implementation of this transformation. The rule $g \in G_R$ has a uniform structure for all stages:

$$g = \langle f_{ap}^g, f_{ir}^g \rangle \quad (47)$$

where f_{ap}^g - the applicability function, which determines whether the rule can be applied to a certain set of input information;

f_{ir}^g - a transformation function that specifies the transformation of the input information.

The rule g set for the transformation $F_g : X \rightarrow Y$ has the form (48).

$$F_g(x) = \begin{cases} f_{ir}^g(x), & f_{ap}^g(x) \\ x, & \neg f_{ap}^g(x) \end{cases} \quad (48)$$

Each function of applicability is a lambda-term of the form [40, 54, 55]:

$$f_{ap} = (\lambda x_1, x_2 \dots x_{n_g}. t_{ap}(x)) a_1, a_2 \dots a_{n_g} = t_{ap}(a_1, a_2 \dots a_{n_g}) \quad (49)$$

where the record λx indicates that this design is λ - term;

x_i - a variable that takes values in the plural $L \cup L^*$;

a_i - argument of the function that sets the value x_i ;
 n_g - number of arguments to be submitted to the input of the conversion function;
 t_{ap} - condition of applicability, a formula that contains n_g variables.

In the general case, the condition of applicability t_{ap} means the existence of a homeomorphism between a directed graph formed by an input sequence of lexemes (as well as syntactic links between them) and a certain reference oriented graph G_{ap} representing a subgraph T_{sn}^e selected by the user of the initial representation of a certain text. As T_{sn}^e can be the initial representation of both the current text T_{sn} and any other text (for example, the thesaurus of the SSA). The condition has a structure (50) and consists of *predicates of identification* [51, 58, 68]. Such predicates make it possible to identify the context of a particular lexeme, and, on the basis of this, conclude that there is no need or need for conversion. Each of the predicates sets a certain condition, and the condition of applicability of the rule is to fulfill all the conditions given by each of the predicates. The number of predicates in the formula sets the number n_g .

$$t_{ap} = c_{p_1}(x_1) \& \dots \& c_{p_n}(x_{n_g}) \& r_{k_{11}}(x_1, x_1) \& \dots \& r_{k_{n_g n_g}}(x_{n_g}, x_{n_g}) \quad (50)$$

Single predicates present in the formula are predicates of identifying lexemes. Such a predicate sets the condition that a certain lexeme (or construct) of the input plural must match. The predicate has a structure (51).

$$c_p(l) = \begin{cases} 1, p = 0 \vee p = l^T \vee p \in P_l \\ 0, p \neq 0 \wedge p \neq l^T \wedge p \notin P_l \end{cases} \quad (51)$$

The work of such a predicate depends on *the template parameter* p . Depending on the type of this parameter, the predicate can be:

- (1) The standard predicate of identification. In such a predicate p - this is a morphological characteristic of lexemes. This predicate determines whether the input lexeme has a given characteristic ($p \in P_l$).
- (2) Predict the identification of keywords. In such predicates p , it is a text representation of the necessary lexeme, and this value is compared with the value of the input ($p = l^T$). This predicate always has a value of 0 for constructs.
- (3) Zero predicate ($p = 0$). This predicate always has a value of 1, regardless of the input lexeme.

A double predicate is a predicate of link identification. This predicate determines whether there is a connection between the two given lexemes of a given type. The predicate has the form (52).

$$r_k(l_1, l_2) = \begin{cases} 1, k = 0 \vee \langle l_1, l_2, k \rangle \in R_{sn} \\ 0, k \neq 0 \wedge \langle l_1, l_2, k \rangle \notin R_{sn} \end{cases} \quad (52)$$

Like predicate lexemes, this predicate has zero modification, which has value 1 regardless of input data. In this case, for the correct use of the condition (50), the condition given by the structure of the lexical analyzer must be fulfilled:

$$k_{ij} = 0, i = j \quad (53)$$

The rule base defines a transformation F that looks like:

$$F_G(L) = \bigcup_{\tilde{L} \in P(L)} \bigcup_{g \in G} F_g^*(\tilde{L}) \quad (54)$$

where $P(L)$ - the set of all subsets L ;

F_g^* - modified function (48), supplemented by additional conditions.

Additional conditions are superimposed on two:

The condition of the order means that all elements of the input subset must be linearly ordered by a certain ratio of strict order G , and has the form:

$$f_{ord}(x) = \begin{cases} 1, \forall x_1, x_2 \in x, x_1 \prec x_2 \vee x_2 \prec x_1 \\ 0, \exists x_1, x_2 \in x, x_1 \not\prec x_2 \wedge x_2 \not\prec x_1 \end{cases} \quad (55)$$

The order of reference that can be used as G :

- (1) Relationship to follow up \prec
- (2) Transitive closure ratio R (43)
- (3) Transitive closure of the relationship given by the links R_{sem}

The condition of consistency f_{ap+}^g determines whether this element $\tilde{L} \in P(L)$ is suitable for processing by the rule g :

$$f_{ap+}^g(x) = \begin{cases} 1, card(x) = n_g \\ 0, card(x) \neq n_g \end{cases} \quad (56)$$

Given these conditions of transformation (48) becomes F_g^* :

$$F_g^*(x) = \begin{cases} f_{ir}^g(x), & f_{ord}(x) \wedge f_{ap+}^g(x) \wedge f_{ap}^g(x) \\ x, & \neg f_{ord}(x) \vee \neg f_{ap+}^g(x) \vee \neg f_{ap}^g(x) \end{cases} \quad (57)$$

Consider the peculiarities of performing various reduction steps within the structure (57).

The transformation of aggregation is a transformation that forms structures (constructs) from linked lexemes, which later become candidates in concepts or contexts of concepts. The peculiarity of the work of transformation is that it applies to the sentence as a whole, that is, to the set of lexemes L associated with a particular sentence S . The transformation forms a plurality of constructs L^* , which in the future is also associated with the sentence S , and is used in conjunction with the initial set of lexemes. All

subsequent transformations are used as the input data of the set, which retains the relation of follow-up \prec , but which in this case is converted into a partial order.

The rules of this transformation differ from the standard structure by the method of application and the method of forming the function of transformation. At this stage, a modified function F_g^* with the form (58) is used.

$$F_g^*(x) = \begin{cases} f_{tr}^g(x), & f_{ord}(x) \wedge f_{ap+}^g(x) \wedge f_{ap}^g(x) \\ \emptyset, & \neg f_{ord}(x) \vee \neg f_{ap+}^g(x) \vee \neg f_{ap}^g(x) \end{cases} \quad (58)$$

The function (58), unlike the function (57), returns an empty set in case of its inapplicability. Due to this, the resulting set does not fall into the elements of the input set L . The function f_{tr}^g is a function of the aggregation of the form (59).

$$f_{tr}^g : \{l_1 \dots l_{n_g}\} \rightarrow \left\{ l^* = \left\langle V^+ \left(\tilde{l}_1^V, \tilde{l}_2^V \dots \tilde{l}_{n_g}^V \right), P_{l_{n_r}}, L_{l^*}^* \right\rangle \right\} \quad (59)$$

where l^* - the resulting construct;

$L_{l^*}^*$ - a set of lexemes or constructs from which the construct is formed;

V^+ - operation of combining numerical representations \tilde{l}_i^V of lexemes of the input plural;

n_r - a specific index ($1 \leq n_r \leq n_g$) that defines which word is the main word in the selected phrase (for example, in a combination of adjectives and nouns, the noun will be the main one). The index n_r determines which lexeme construct l^* imposes its morphological characteristics P_{l^*} . In the case that the construct is formed not by words (for example, date), this parameter has no content and can be considered a zero.

On the basis of this index, the operation M of finding the word of the phrase is constructed:

$$M(L) = l_{n_r} \quad (60)$$

Actually, the transformation of aggregation is performed by recursive application of the function (58) according to the scheme (61).

$$F_{ag}(\tilde{L}) = \begin{cases} \tilde{L}, & F_G(L \cup \tilde{L}) = \tilde{L} \\ F_{ag}(F_G(L \cup \tilde{L})), & F_G(L \cup \tilde{L}) \neq \tilde{L} \end{cases} \quad (61)$$

Transformation of concept identification is a transformation that forms a set of concepts X based on a plurality $L \cup L^*$. This transformation is performed after the conversion of aggregation and uses the constructions L^* generated by this transformation.

In general, this transformation has a standard structure. The applicability condition often has a trivial structure:

$$t_{ap} = c_{p_1}(x_1) \quad (62)$$

Where c_{p_1} identifies nouns. In this way, the phrases, the main word for which is the noun, and for which the constructs were formed during the aggregation stage, immediately fall under the rule, and from them the concepts are formed. The transformation function f_{ir}^g is trivial:

$$f_{ir}^g(L^*) = x(l_1^*) \in X \quad (63)$$

where $x(l)$ - the function of forming a concept x from lexeme l .

At the same time, the lexeme L form those concepts that are described in one word, and form the constructs L^* that are described in a few words.

Transformation of link identification is a transformation that forms a set of semantic relationships between concepts R_{sem} , semantic relationships between concepts and lexemes/constructs R_{sem}^* on the basis of sets $L \cup L^*$ and X .

Transformation has a standard structure, with the condition of applicability as a rule has the form:

$$t_{ap} = c_{p_1^*}(x_1) \& c_{p_1}(x_2) \& \dots \& c_{p_{n_g-2}}(x_{n_g-1}) \\ \& c_{p_2^*}(x_{n_g}) \& r_{k_{12}}(x_1, x_2) \& \dots \& r_{k_{n_g}}(x_{n-1}, x_{n_g}) \quad (64)$$

where p_1^*, p_2^* - conditions that determine the type of lexemes or constructs between which (or their corresponding concepts) will be established a connection;
 $p_1 \dots p_{n_g-2}$ - conditions defining keywords that are specific to the type of communication identified by the rule.

The transformation function has the form (65).

$$f_{ir}^g(L^*) = \begin{cases} \langle x(l_1^*), x(l_{n_g}^*), k_{sem}^g \rangle \in R_{sem}, & x(l_1^*) \in X \wedge x(l_{n_g}^*) \in X \\ \langle x(l_1^*), l_{n_g}^*, k_{sem}^g \rangle \in R_{sem}^*, & x(l_1^*) \in X \wedge x(l_{n_g}^*) \notin X \\ \langle x(l_{n_g}^*), l_1^*, k_{sem}^g \rangle \in R_{sem}^*, & x(l_1^*) \notin X \wedge x(l_{n_g}^*) \in X \\ 0, & x(l_1^*) \notin X \wedge x(l_{n_g}^*) \notin X \end{cases} \quad (65)$$

where $l_1^* \dots l_{n_g}^*$ - the constructs of the input plural;

$x(l)$ - the function of forming a concept x from a lexeme l ;

k_{sem}^g - The type of semantic relation identified by the current rule.

This transformation determines which communication is identified in its result. If at the stage of identifying concepts, ones have been identified and concepts created on both the basis l_1^* and on the basis $l_{n_g}^*$, then the formed link will be assigned to the plural

R_{sem} . If only for one of the constructs the concept was not created, then such link would fall into the set R_{sem}^* and will be used to identify the contexts of the concept associated with it. If the concept was not created on the basis of any of the concepts, then the connection is not created. Most likely, such situation means that there is no rule of aggregation that would combine these constructs into one.

Transformation of context identification is intended for the analysis of the contexts of concepts given by relations $R_{sem}^* \cup R_{sem}$, and the formation of attributes based on them, which then define the functions of interpretation of concepts.

Transformation of context identification uses a modified link identification predicate (52), which is intended to identify semantic relationships between concepts. The predicate has the following form:

$$r_k^*(x_1, x_2) = \begin{cases} 1, & \langle x_1, x_2, k \rangle \in R_{sem}^* \cup R_{sem} \\ 0, & \langle x_1, x_2, k \rangle \notin R_{sem}^* \cup R_{sem} \end{cases} \quad (66)$$

where x_1, x_2 - certain concepts $x_1, x_2 \in X$.

The applicability function does not use predicate lexemes identification (51), and has the following form:

$$t_{ap} = r_{k_{12}}^*(x_1, x_2) \quad (67)$$

The transformation function for this case is also trivial and has the form:

$$f_{tr}^g(L^*) = a_{l_2}^g(x(I_1^*)) \in A_X \quad (68)$$

where $a_l(x)$ - the function of forming an attribute with a lexeme or construct l for the concept x .

The transformation of the previous structuring F_{ps} is not part of the reduction operator, but is carried out in the same way as its parts. This transformation can complement the parsing process.

Pre-structuring of text is often necessary when working with weakly structured documents. This is especially important when processing tables that contain headers, but the values of table cells contain a natural language that needs to be analyzed in the same way as the NL text to obtain the necessary information. Also, as a rule, concepts are contained in the headings of the text sections, and the concept in the header may be the name of a particular category, which includes the concepts given in this section.

All that needs to be done with the previous structuring is to identify the named areas in the text (such as sections or columns of the table). This requires:

- (1) Define lexemes of the text that form the name of the named area.
- (2) Define lexemes of the text that form the found named domain.

The name of an area is, as a rule, a certain block of text (often sentence), which is highlighted in a markup in a special way, which, in turn, can be accomplished in two ways:

- (1) Explicit - using a specific separator character, such as a comma in CSV (Comma Separated Values) files.
- (2) Implicit - using the display settings of the text, for example, by selecting another font.

When reading documents that support implicit allocation, it is necessary to perform additional processing in order to obtain metadata for displaying lexemes. In this case, the set of signs P_l of a lexeme l will consist of two parts - morphological signs P_l^{mph} and metadata of reflection P_l^{dis} .

$$P_l = P_l^{mph} \cup P_l^{dis} \quad (69)$$

The function of the previous structuring transformation has a standard look (48)–(52). However, it uses the specialized structure of the applicability condition (50) having the form (70).

$$t_{ap} = c_d(x_1) \& c_t(x_2) \& \dots \& c_t(x_{n_g-1}) \& c_d(x_{n_g}) \& r_0(x_1, x_2) \& \dots \& r_0(x_{n-1}, x_{n_g}) \quad (70)$$

The condition uses only two template parameters: d - constraint condition, and t - condition comparison:

- (1) For an explicit way of defining, c_d identifies a specific separator character, while c_t does not impose any conditions ($c_t = c_0$).
- (2) For an implicit method of assignment c_d , identifies any lexemes that, according to the styles of the current document, do not relate to the headings, whereas c_t on the contrary, it identifies any lexemes that belong to them.

Syntactic relationships between lexemes are not used at the stage of previous structuring, therefore the conditions on them are not superimposed ($\forall i, \forall j, k_{ij} = 0$).

If you find the lexemes belonging to the named block, in the explicit method of specifying the rule is similar, whereas when implicitly the conditions are changed places.

The function of the transformation of the previous structuring performs a transformation of the form (71) or (72), where p_n and p_n^* - the property of the membership of the lexeme to the named area and the affiliation of the lexemes to the name of the named area, respectively.

$$f_{ir} : \langle l^T, P \rangle \rightarrow \langle l^T, P \cup \{p_n\} \rangle \quad (71)$$

$$f_{ir} : \langle l^T, P \rangle \rightarrow \langle l^T, P \cup \{p_n^*\} \rangle \quad (72)$$

Conditions p_n and p_n^* , if necessary, can be used at the next steps by setting the appropriate parameters in the rules templates. In particular, you can include text headers in the source ontology as objects.

3.5 The Structure of the Reduction Operator and Its Components

Whereas the principle of the work of each of the predicates included in the applicability condition (50) depends solely on the associated template parameter p or k , then any applicability function is identified by a set of such parameters. The conversion function f_{ir} , on the other hand, depends solely on the executable step. As a result, the conversion rule g can be represented in a much more compact form - in the form of a rule template \tilde{g} that appears as a formula of the form (73).

$$\tilde{g} = \langle \{p_i\}, \{k_{ij}\} \rangle \quad (73)$$

In a trivial manner, you can construct the transformation of the interpretation of patterns:

$$F_{int} : \tilde{G} \rightarrow G \quad (74)$$

The formula (73) is much more convenient for the user, as it is easier to represent it in a textual form. However, the process of forming conversion rules can be simplified even more.

The formula (73) imposes a certain condition on the topology of the graph structure formed by the input plural and the connections between its elements. Whereas both constructs and concepts within the framework of operations performed on them can be considered as an analogue of lexemes, then the graph structure that they have formed can be regarded as an analogue of the structure formed by lexemes, namely, the primary structure of the text (13). An important consequence of this fact is that to create templates of the form (73) it is easy to develop a procedure for automated creation, because such a procedure is reduced to a simple implementation of the subgraph selection function $G_{ap} \subset T_{sn}^e$. It is important to note that as T_{sn}^e can be used the primary structure of the processed text, that is, the formation of templates can act as an additional step after parsing and before aggregation (25).

Another important consequence of the existence of a homomorphism between the primary structure of the text and the structure of information from which it is allocated is the possibility of constructing a specialized NS NS' intended for work not with objects of ontology, but with lexemes of the primary structure [58]. An interactive document built on its basis can be used to form G_{ap} .

The procedure for creating templates can be simplified even more, creating specialized reference texts that contain only examples of the use of some verbal constructions. In this case, you can exclude the step of selecting a subgraph by accepting $G_{ap} = T_{sn}^e$.

4 Formation of Ontological GIS Applications and Transdisciplinary Representation of Geospatial Information with Their Help

As already mentioned, experts in the corporate information environment can be represented as a system {action \rightarrow results} and represent the analogue of the natural system SN . Based on this system, you can build interactive document types (6). To create an interactive document, you can build a specialized transformation of the form (75).

$$\check{G} : O \Rightarrow SN \quad (75)$$

Consider the natural systems SN characterized by a set n of “actions” $\check{x}^1 \dots \check{x}^n$ and one “result” \check{y} [57], which are connected by dependence:

$$\check{y} = \check{f}(\check{x}^1 \dots \check{x}^n) \quad (76)$$

When forming a natural system based ontologies dependence \check{f} in the general case, has the following form:

$$\check{f}(\check{x}^1 \dots \check{x}^n) = D\left(Q_n\left(Q_{n-1}\left(\dots Q_1\left(Q_o(X), \check{x}^1\right), \dots, \check{x}^{n-1}\right), \check{x}^n\right)\right) \quad (77)$$

where X - the set of objects of the initial ontology O ;

Q_i - auxiliary processing functions;

D - a mapping function that allows the user to display the results of processing objects.

You can define a basic set of auxiliary functions as follows:

Hierarchical Filtration Function (78).

$$Q_h(X, x^*) = \{\bar{x} \in X | \bar{x} R x^* \vee (\exists \tilde{x}, \tilde{x} \bar{R} x^* \wedge \bar{x} \in Q_h(\tilde{x}))\} \quad (78)$$

where X - the set of objects belonging to a certain ontology;

x^* - the object in respect of which the filtration is carried out;

\bar{R} - a certain relation between objects.

This feature allows you to filter objects that are related to a particular relationship \bar{R} . Also, the condition is verified recursively, which is equivalent to the use of a transient circuit \bar{R} .

The ratio \bar{R} should be based on the set of links between the objects R :

$$\begin{aligned} \exists k, \langle x_1, x_2, k \rangle \in R &\Rightarrow x_1 \bar{R} x_2 \\ \exists k, \langle x_1, x_2, k \rangle \in R &\Rightarrow x_2 \bar{R} x_1 \\ \exists k, \langle x_1, x_2, k \rangle \in R \vee \langle x_2, x_1, k \rangle \in R &\Rightarrow x_1 \bar{R} x_2 \end{aligned}$$

where x_1, x_2 - certain objects;
 R - a plurality of links between objects;
 k - the type of relation between objects.

If necessary, the type of relation k can be fixed, which will result in the formation of the relation \bar{R}_k given by the formulae (79)–(81).

$$\langle x_1, x_2, k \rangle \in R \Rightarrow x_1 \bar{R}_k x_2 \quad (79)$$

$$\langle x_1, x_2, k \rangle \in R \Rightarrow x_2 \bar{R}_k x_1 \quad (80)$$

$$\langle x_1, x_2, k \rangle \in R \vee \langle x_2, x_1, k \rangle \in R \Rightarrow x_1 \bar{R}_k x_2 \quad (81)$$

The attribute filtering function looks like:

$$Q_a(X, A) = \{\bar{x} \in X | A \cap A_{\bar{x}} = A\} \quad (82)$$

where X - the set of objects belonging to a certain ontology;
 A - the set of attributes for which filtration is performed;
 A_x - the set of attributes of the object x .

The function Q_a allows you to select a set of objects that have a specific attribute or specific value of a particular attribute.

The context-sensitive function allows you to establish links between objects described in documents belonging to a specific information environment. This function is a combination of two functions.

The indexing function looks like:

$$Q_I(C) = \bigcup_{T \in C} \bigcup_{l \in L_T} \{ \langle V(l), V(T) \rangle \} \quad (83)$$

where C - the input set of documents;
 T - a certain document;
 L_T - a set of lexemes that forms a textual representation T ;
 $V(l), V(T)$ - lexemes l and document T identifiers respectively.

The indexing function results are used by the search function:

$$Q_S(I, l) = \{T | \langle V(l), V(T) \rangle \in I\} \quad (84)$$

where I - index, which is the result of work Q_I ;
 $V(I), V(T)$ - lexemes I and document T identifiers respectively.

With the help of the search function, you can form relationships between objects that belong to different ontologies (or between an ontology object and an unstructured information resource). It is necessary to take into account two features:

- (1) The object context usually consists of a large number of attributes, which, when constructing the index, can be regarded as a single text.
- (2) Both the name and the context of the object in the general case consists of many lexemes.

Taking into account these features, the function of the context link will look like:

$$Q_c(x) = \bigcup_{l \in L_x} Q_S(Q_I(C), l) \quad (85)$$

where C - a set of documents, representing the information environment, within which a bunch is carried out;

x - the object with which the bunch is made;

L_x - textual representation of the context.

l - a certain lexeme.

The main disadvantage of the function (85) is that it forms an unordered set of documents as a result, which can be quite large in size, and therefore uncomfortable for processing by an expert. There are two ways to solve this problem:

- (1) Assign on the set of results the order of the relation, which can be the relevance ratio R_{rel} :

$$T_1 R_{rel} T_2 \Rightarrow \text{card}(L_{T_1} \cap L_x) > \text{card}(L_{T_2} \cap L_x) \quad (86)$$

where L_{T_1}, L_{T_2} - text representations of documents T_1, T_2 ;

L_x - The context of the object x with which the contextual link was implemented.

- (2) Delete from the result documents with insufficient relevance that can be used in the function (85) of the intersection operation instead of the merger operation.

5 Transdisciplinary Presentation of Information Through Interactive Documents

The transdisciplinary representation of sets of ontologies is based on the transformation between two hypersets:

$$f^{ct} : R \rightarrow F \quad (87)$$

where R, F - the set of links and functions of the interpretation of a certain ontology O .

For a set of ontologies, you can construct a hyperset:

$$\mathfrak{R} = \bigcup_i R_i, F = \bigcup_i F_i \quad (88)$$

where i - the index defining a certain ontology $O_i = \langle X_i, R_i, F_i \rangle$.

On these hypersets it is possible to construct a converse transformation to f^{ct}

$$f^{tt} : F \rightarrow \mathfrak{R} \quad (89)$$

That is, a certain connection between objects of various ontology themes can be represented by a non-empty set of interpretive functions from data ontologies.

This transformation can be extended to a set of unstructured texts, applying to each of the texts the reduction operator (24).

Let's consider the procedure of transdisciplinary representation of a certain set of ontologies by means of interactive documents. This representation is based on the function (85). To implement it, you need to execute transdisciplinary transformation using the function:

$$Q_{TI}(C) = \bigcup_{x \in X_C} \{Q_C(x)\} \quad (90)$$

where C - the set of ontologies;

X_C - a set of objects belonging to the association of ontologies $\bigcup_{O \in C} O$.

The application of the function $Q_C(x)$ generates a set of objects from different ontologies, representing a certain hyperratio between the corresponding objects. With the help of the hyperratios thus formed, a transdisciplinary representation O' of the set of ontologies $C = \{ \langle X_i, R_i, F_i \rangle \}$ can be constructed:

$$C \xrightarrow{Q_{TI}} \langle \bigcup_i X_i, \bigcup_i R_i \cup Q_{TI}(C), \bigcup_i F_i \rangle \quad (91)$$

Transform (91) sets the most complete representation of the available in C information, which is not always convenient. Often it is necessary to perform a representation of one selected ontology O . In this case (90) it is necessary to change:

$$Q_{TO}(O, C) = \bigcup_{x \in X_O} \{Q_C(x)\} \quad (92)$$

where C - the set of ontologies;

X_O - a set of objects belonging to ontology O .

For the construction of systems that use the transdisciplinary representation of information, the following statements are important:

Statement 2.1. On the basis of one ontology O belonging to a plurality of documents C , it is possible to form an arbitrary number of NS.

Argument. Let there exist a certain type of NS (76), built on the basis of ontology $O = \langle X, R, F \rangle$ with the help of transformation (75), and a set of documents C . Whereas the NS is determined by its function of the form (77), it is possible to form an arbitrary number of natural systems, changing the composition and the number of functions $Q_0 \dots Q_n$. In particular:

- (1) If you lock an object $\tilde{x} \in X$ arbitrarily and apply the hierarchical filtering function to the ontology, then you can take $Q_0(X) = Q_h(X, \tilde{x})$ and receive a system that represents a certain class of objects;
- (2) If we select as a Q_0 function that performs the transdisciplinary representation (91), then we obtain a system that reflects the transdisciplinary representation of the ontology O' .

Consistently applying the function Q_h with different parameters and functions performed transdisciplinary representation, you can create any number of initial ontology O changes, the use of which in the formula (77) and will create as many natural systems.

The consequence of Statement 2.1 is that on the basis of one ontology, it is possible to construct an arbitrary number of interactive documents.

Natural systems of the species (76) are the simplest variant of natural systems. However, such system may be ineffective, in particular in the following cases:

If the information can be presented in several ways (for example, in the form of a table and an e-card), then it is usually not premature to know which method is more convenient for an expert;

If the information is different in content and structure, then it may be necessary to represent its various subsets in different ways;

So often there is a need to form the content from another ontology several natural systems act independently and accept the same set of input “action” from the user, and giving it a different “results” from which he can choose the one that meets the challenge.

Statement 2.2. A combination of independent natural systems SN_i that accept the same set of “actions” is a natural system.

Argument. According to the definition of the natural system in [57], if at least one element in the system during the action will give the same result, as in the work independently of other elements of the system. If we consider each of SN_i as an element of a particular metasystem S , then due to their independence SN_i , their work in the system will not be different from work outside the system. Therefore, S will also be a natural system.

Such system S will be an extension (76) and will have the form (93).

$$\begin{aligned} \check{y}^1 &= \check{f}^1 \left(\check{x}^1 \dots \check{x}^n \right) \\ &\vdots \\ \check{y}^m &= \check{f}^m \left(\check{x}^1 \dots \check{x}^n \right) \end{aligned} \tag{93}$$

6 Transdisciplinary Representation of Geospatial Information with Ontological GIS Applications

The ontological GIS-application of the form (7) is also formed on the basis of the NS, which can be represented by the formula (76). In the simplest case, such NS can be regarded as a normal NS with a special display function - the function of mapping as an electronic map marker D_m . However, the main feature of this function is that not all concepts can be represented as a marker, but only those that are elements of the layer (which, for example, can be represented by a certain category of concepts) and whose attributes contain geographic information. Therefore before using D_m it is necessary to filter:

$$\tilde{T}_{lr} = Q_a(Q_h(X, x_{lr}), A_{geo}) \quad (94)$$

where \tilde{T}_{lr} - a taxonomy formed by a subset of ontology concepts that can be represented as markers of a certain layer;
 x_{lr} - an object that defines a class of objects that forms a layer of GIS;
 A_{geo} - A set of attributes that can contain geographic information.

As a result of this filtering, you can get a set of taxonomies (usually non-intersecting). This set can be displayed with D_m :

$$D_m\left(\bigcup_{x_{lr} \in X_{lr}} \tilde{T}_{lr}\right) \quad (95)$$

where X_{lr} - the set of classes of concepts representing layers of GIS;
 x_{lr} - a concept that defines the class of concepts that form a layer of GIS;
 \tilde{T}_{lr} - taxonomic representation of the layer x_{lr} given;
 D_m - function of displaying an object as an e-card marker.

It is formed on the basis of such an NS procedure and will form an ontological GIS-application of the form (7). Bring the formula (7) to the standard structure of the interactive document can be as follows:

A set of geographic entities X_g can be formed by combining taxonomic representations of GIS layers (94), as in the formula (95):

$$X_g = \bigcup_{x_{lr} \in X_{lr}} \tilde{T}_{lr} \quad (96)$$

On the basis of the set X_g , you can also form a set $R_g \subset R$ with the formula (97). Typically, such set will correspond to the union of sets of links of taxonomic

representations \tilde{T}_{lr} of GIS layers, however, and specialized links between objects of different taxonomic representations are possible.

$$R_g = \{ \langle x_1, x_2 \rangle \in R \mid x_1 \in X_g \wedge x_2 \in X_g \} \tag{97}$$

The set of states S appears as a sequence of “results” \tilde{y} of the work of the natural system, which were formed in response to the “actions” provided by the user:

$$S = \{ \tilde{y} \} \tag{98}$$

The set of analytic operations over geographic objects A^{op} essentially represents a subset of the set of auxiliary functions Q forming the dependence (77):

$$A^{op} \subset Q \tag{99}$$

This set can be divided into two subsets: the set of operations of the analysis of the spatial relation of objects A^{rel} and the set of measuring operations A^{ms} .

Table 1. Operations of the analysis of spatial relations of objects

Binary relations	Inclusion ratio	Matches
		Contains in itself
		Contained in
	Intersection ratio	Intersects with
		Borders on
		Separated from
Relationship of adjacency	Neighborhood	
Other relations	Distances	

The operations of the analysis of the spatial ratio of objects allow the selection of a plurality of objects, such that the geospatial information associated with them meets a certain criterion. Such operations can be used to filter objects, in combination with the functions (78)–(84). The classification of such operations is presented in Table 1.

Most of these relationships are binary. When the user specifies a geospatial object-standard (point or polygon), it is necessary to highlight among objects of ontology such ones that contain geospatial information, which in turn is linked to the given binary relation with the standard:

$$Q^r(X, g) = \{ x \in X \mid \langle A_{geo}(x), g \rangle \in r \} \tag{100}$$

where r - a certain binary geospatial relation;

X - a set of taxonomy objects;

g - spatial object acting as a reference;

$A_{geo}(x)$ - a function of the sample of the geospatial information associated with the object x .

The distance ratio can also be used for filtering if you set a specific maximum distance value d_{max} . This function is extremely useful as an alternative to the “coincides” relation for cases where both the reference and the set of geospatial information for which filtration is performed are points (for example, when handling a click-by-click event). This function can be represented as follows:

$$Q^d(X, g) = \{x \in X | d(A_{geo}(x), g) < d_{max}\} \quad (101)$$

where d - function of distance;

X - a set of taxonomy objects;

g - spatial object acting as a reference;

$A_{geo}(x)$ - function of the sample of geospatial information connected with the object x ;

d_{max} - the value of the maximum permissible distance.

Measuring operations A^{ms} are fundamentally different from the operations of the analysis of spatial relations of objects. Their result is a certain numeric value, for presentation of which the user needs a special display function D_n . The classification of such operations is given in Table 2.

The main difference between computational operations is that they accept as an argument not a standard for filtering, but a particular object over which geospatial information is bound to perform the calculation. However, if necessary, on their basis, one can construct the functions of filtration, in particular, similar to (101).

We apply all the elements of the set of ontologies C to the following conditions: each of them has a non-empty subset of objects containing geographic information:

$$\forall O \in C, Q_a(X_O, A_{geo}) \neq \emptyset \quad (102)$$

where X_O - the set of objects ontology O ;

A_{geo} - a set of geographical attributes;

Q_a - attribute filtering function (82).

An arbitrary GIS can be represented as follows [13, 61, 66, 67]:

$$\langle X_g, R_g, A^{op}, T^s \rangle \quad (103)$$

where X_g - the set of geographical entities over which analytical operations are performed to solve the problem;

R_g - a set of relationships between geographic objects that determine the type of executed operations;

A^{op} - a set of analytical operations on geographic objects performed in the process of solving a problem;

T^s - the set of states of the task, which are visualized on the map in the process of its solution;

Table 2. Measuring operations

Define the coordinates of the point on the map	Determination of the coordinates of the point of intersection of two straight lines	
	Definition of coordinates of centroids and centers	
Measurement of distances	The length of the line/curve between two points	
	The shortest distance for the shortest straight line, taking into account the sphericity of the earth's surface	
	The shortest distance from the given point to the landfill and between the polygons	
	Functional distance on a given route using turning points/taking into account the uneven terrain ("on the ground")	
	Distance and cost over the network/over the surface	
Measurement of polygons	Perimeter	
	Area	
	Forms	Spatial integrity
		Ratio of large and small axes
		Perimeter ratio to square
The development of boundaries		
Measuring volumes		

Statement 2.3. Each GIS can be represented as a definite natural system.

This statement follows from the definition of the natural system [57], because GIS can be represented as a system of n influences and one "result" (76), which serves as the representation of the user geospatial information in the right form.

Statement 2.4. The set of all GIS can be represented as a hyperset of natural systems.

Functional characteristics of GIS can be interpreted on the basis of affine space. The affinity space above the field K is called three:

$$\langle A_n, V_n, + \rangle \tag{104}$$

where A_n - the set, the elements of which are called points;

V_n - vector space above the field K ;

$+$ - a binary operation $A_n \times V_n \rightarrow A_n$ that satisfies certain axioms [69].

A set of points A_n is called n -dimensional affine space over K , and the vector space V_n is a guide for A_n [69, 70].

Such space can be easily built as part of a certain ontological GIS application. To do this, you need to use the sample function of the geospatial information associated with the object $A_{geo}(x)$:

$$A_{geo}^x = \{A_{geo}(x) | x \in X\} \quad (105)$$

The set of points $A_n = A_{geo}^x$ formed in this way allows form affine space over a field \mathbb{R}^2 or \mathbb{R}^3 (depending on the structure of the information in the ontological GIS application initialization and the requirements of the solvable problem). This operation is carried out by specialized transformations:

$$\tilde{G}_A : \langle X, R, F \rangle \rightarrow \langle A_{geo}^x, V_n, + \rangle \quad (106)$$

where X, R, F - the set of objects, connections and functions of the interpretation of ontology;

A_{geo}^x - formed on the basis of X (105) the set of geographic points;

$V_n, +$ - vector space over a field \mathbb{R}^2 or \mathbb{R}^3 a certain binary operation.

Points belonging to a plurality of geographic information A_{geo}^x will be related to the vectors that belongs to V_n as follows:

- (1) For each ordered pair of points A, B , you can put in a matching vector $\overrightarrow{AB} = v \in V_n$ that can be labeled as $v = B - A$.
- (2) For a given point A and a vector $v \in V_n \exists! B \in A_n : B - A = v$.
- (3) For arbitrary three points A, B, C , the identity of the triangle is fulfilled in V_n :

$$(B-A) + (A-C) + (C-B) = 0 \quad (107)$$

You can select properties:

- (1) The two identical points correspond to the zero vector.
- (2) A vector determined by a pair of points B, A opposite to a vector determined by a pair of points A, B .
- (3) If $B - A = B - A'$ so $A = A'$.
- (4) If we choose a certain starting point O , then $\forall v \in V_n$ only one point $B \in A_n$ corresponds, and $\forall B \in A_n$ corresponds only one vector $v = A - O$.

Consider the plane, as affine space A_2 , and consider all mutually $A_2 \rightarrow A_2$ unambiguous mappings that keep the distances between the points and the angles between the vectors. The classification of the main such mappings is shown in Table 3.

Each of the main types of mapping corresponds to a certain function that performs such a reflection over a certain affine space \tilde{A}_n formed by a certain subset of ontology $\tilde{X} \subset X$ objects. These functions have a structure (108).

$$Q_t(O) = \tilde{G}_A^{-1} \left(t \left(\tilde{G}_A(O) \right) \right) \tag{108}$$

where $\tilde{G}_A, \tilde{G}_A^{-1}$ - the transformation (106) and the inverse of it;
 t - a certain affinity mapping $A_n \rightarrow A_n$.

Functions Q_t, Q_{rt}, Q_{in} and others specified by specialized views can be used in combination with the standard functions of an interactive document for building ontological GIS applications [71–74]. Ontological GIS applications, in turn, provide transdisciplinary integration of geospatial information through interactive documents and reflections of the affinity space.

Table 3. Classification of main reflections of the affinity space

Group	Marking	Note	F-n Int. document
A group of all non-degenerate linear mappings $V_n \rightarrow V_n$	$GL_n(R)$, general linear group	Specifies all mutually unambiguous vector space transformations V_n	
Linear display group $V_n \rightarrow V_n$, having matrices with determinant 1	$SL_n(R)$, special linear group	Mutually unambiguous transformations V_n , that store lengths and volumes	
Broadcast group	$T_n(R)$	Parallel space transfers A_n	Q_t
Orthogonal group O_n	$O_n(R)$	Mutually unambiguous transformations of the real space $V_n \rightarrow V_n$, that keeping corners and distances	
Group of turns	$SO_n(R)$, special orthogonal group	Keeps angles, distances, and orientation of vector systems	Q_{rt}
The set of inversions		Does not form a group, since the product of two inversions will have a matrix with a determinant 1, and therefore, represents a certain rotor	Q_{in}

7 The Practical Approach for the Taras Shevchenko Portal

The reflection of the transdisciplinary representation of geospatial information is presented on the example of the ontological GIS formation, as a certain means of reflecting the life of Taras Shevchenko (see Fig. 1).

The basis of ontological GIS is an interactive document for which semantic hyperratio and categories of classes must be defined, which collectively provide processes for the formation of transdisciplinary interactive mappings of complex research

topics. It is necessary to determine their semantic connectivity, which is formed on the basis of the use of interdisciplinary relationships that provide structuring and actualization of information, meaningful interactive activity of the researcher of the work of Shevchenko (Table 4).

Table 4. Categories of transdisciplinary representation on the basis of hyperratios

<i>I. Ontological categories of the upper level</i>	
Types of abstractions and hyperratios	Classification, generalization, aggregation and association
Principles of classification	Fundamental principles of Aristotle’s dichotomy, Pearce Trichotomy and Leibniz lattices, natural classification
Types of classification division	Taxonomic, logical, spatial and temporal
Inheritance of signs	Multiple
<i>II. Language-ontological picture of the world</i>	
A. Level of concepts	
Concept-role relation	Two representativeness of the primitive are produced, respectively named as joining and correlative communication. The natural language is translated into different grammatical forms of verbs to <i>have</i> and to <i>be</i>
Semantic relations	Classification - genus-kind, integral-part, class-element, above-located-below-located, class-subclass; Signs - object-attribute, object-action; Quantitative - have a measure
Linguistic relations	Hyponymy, metonymy, synonymy and antonymy
B. Level of primitives	
Predictive relations	Generalized-unitary character, separately-unitary and individual-unitary;
Semantic-role relations	Classification - to have a name, to be a benchmark; Signs - to have an attribute value, to have a meaningful mode of action; Quantitative - have a significant degree; Comparison - equal, comparative, greater, greater or equal, less, less or equal, incomparable; Accessories - simple accessories; Simple times - to be at the same time, to be earlier, to be later, to coincide in time, to intersect in time, to be inward in time, to begin simultaneously, to end in time; Simple spatial - to coincide in space, to be on the left, to be on the right, to be ahead, to be behind, to be from above, to be from below, to be steer, to intersect in space, to touch, to be placed on, to be placed in; Causal - to be a goal, to be a motive, a cause - a consequence; Instrumental - to serve for, to be a tool for, to promote, to be a tool, to be an auxiliary means; Informational - to be the sender, to be the recipient, to be the source of information; Ordinal - to be next, to be current, to be closest; Modalities - possibilities, actual incarnation, necessity; modification; Quantification - universality, existence, uniqueness, identity, opportunity, necessity; Correlation

The above categories quite fully provide the use of the mechanism of recursive reduction in the formation of interactive documents and display of geospatial information.

These mechanisms provide the construction of different classes of taxonomies, which, in turn, can be identified or expanded in such a way that the construction of vertex names in each taxonomy, from the notion of the lower level to the notion of the upper level, forms a logical chain. For this we consider all the connected vertices of the graph as certain tautologies, which are constructed according to the rule: <concept> - property-relation - <concept>. The vertices of graphs are formed by certain, content-filled concepts, and the graphs can be defined as tautologies, taking into account certain limitations of intuitive understanding of concepts. Formulae, which are formed from all vertices, having an attitude relation, can be regarded as the set of all admissible tautologies and consist of allegations of the affiliation of concepts to a particular taxonomy. The tautologies are formed on the basis of the thematic linking of the concepts of classes bearing the names of the above taxonomies.

Taxonomies in the ontological GIS environment that provide grouping classes of ontology objects of the research area correspond to thematic layers of the electronic map, and just the objects belonging to the corresponding class corresponds to objects of the thematic layer. The taxonomy of ontology objects, which corresponds to the legend of the map, is formed on the basis of establishing relationships between concepts and classes, for example, the relation "part - the whole".

The representation of the taxonomy classes in the form of thematic layers of the map allows us to combine the notion of place and time with the concepts of facts and events in an unknown to this combination, from a new angle of view.

All categories, taxonomies and classes are formed on the basis of a certain classification. This classification may be changed. The entire hierarchy, names and properties of each category and class can be edited at any time and under any conditions. It is also possible to expand their list. That is, the ontology of such a complex subject of research as the life and work of Taras Shevchenko is dynamic and provides as a broad and rather deep reflection of all the events and facts that have come to pass in his life, and those that have existed to the present.

Application of the function of membership allows you to define classes of taxonomies, concepts-notions that can be represented by thematic layers of the map. Also, concepts that directly form taxonomies can be reflected in the legends of these thematic layers.

The legend of the map includes thematic layers, similar to the name taxonomies, as classes of concepts that have the same set of properties and objects of the layers. The correspondence between thematic layers of maps and taxonomies reflecting the structure of information arrays is realized on the basis of determining the level of similarity of fuzzy formal concepts of the investigated subjects.

Description of the concepts of the subject area in the form of certain objects on the map is limited to the fields of attribute information, and the service of attachments allows you to attach only the information that is physically available to the user. Due to the combination of different types of databases, in taxonomy, attributes of objects can be represented not only in table form, but also in text, as well as in the form of thematic hyperlinks to distributed information resources in the network (see Figs. 2 and 3).

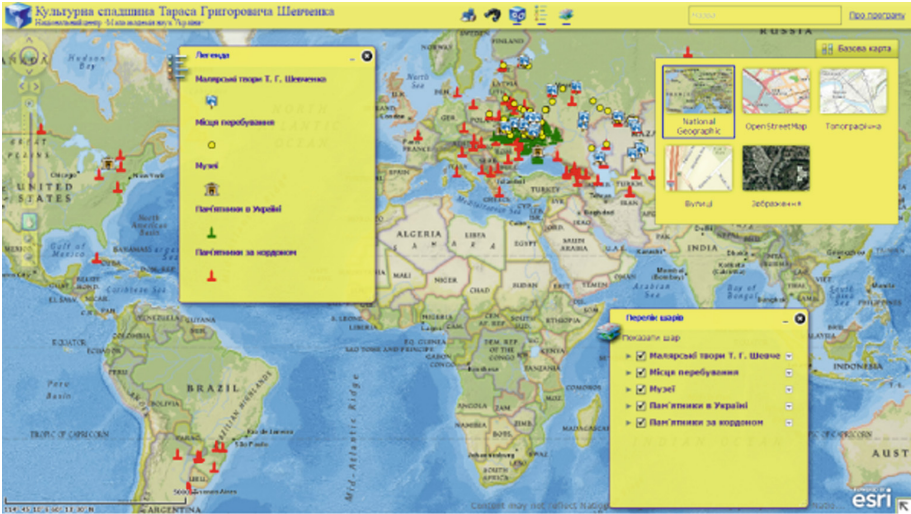


Fig. 2. Thematic layers on the map dedicated to the life and work of Taras Shevchenko

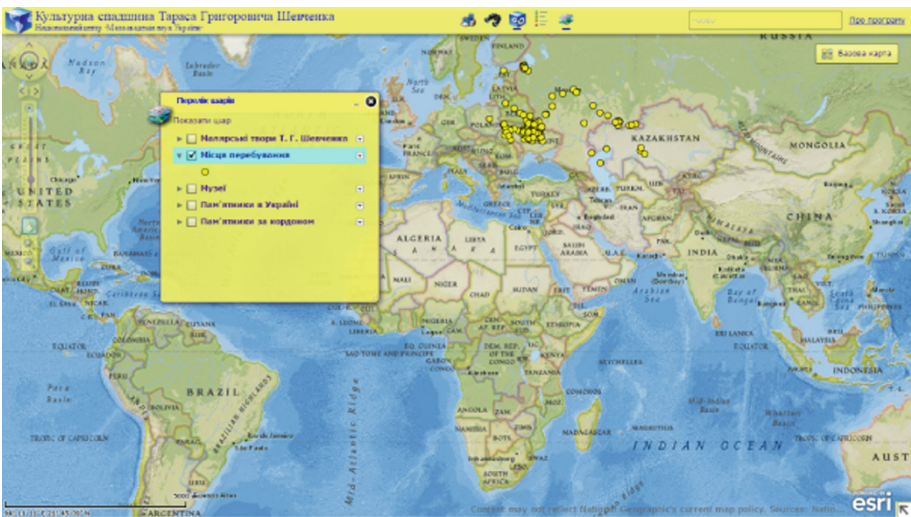


Fig. 3. Objects of the thematic layer “Places of stay

The mechanism of recursive reduction provides a dynamic change in thematic profiles of research in the subject area. Interactive documents that are formed at the stage of semantic analysis of input documents (see Fig. 1) provide such a dynamic change based on the use of hyper properties (see Table 1).

An example of their use is shown in Fig. 4 - a dynamic transition from places of residence to the location of archival documents that reflect the various events of the life

of Shevchenko. Ontological GIS tools provide the opening of an information card window containing information about its official name.

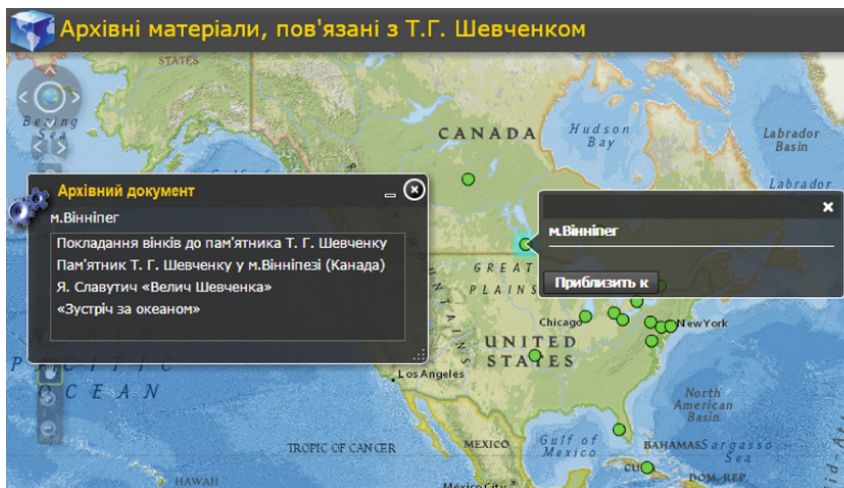


Fig. 4. Window of the “Archival Document” widget

The formation of the indicated mapping is realized on the basis of the mechanism of recursive reduction in the processing of a certain object of the ontological register of archival documents related to the life, creativity and honoring the memory of Taras Shevchenko in the ontological interface (see Fig. 5).

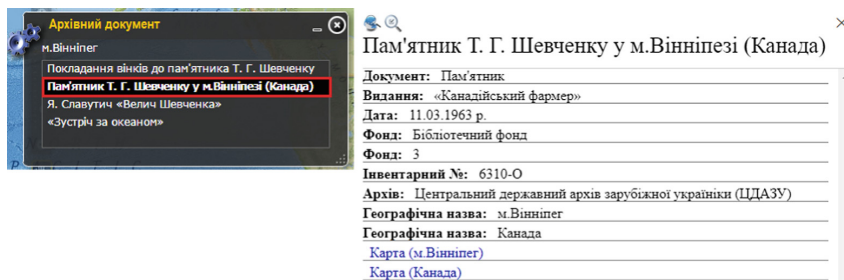


Fig. 5. The transition between the geographic information system and the ontological interface.

Thus, the interactivity of documents formed on the basis of transdisciplinary representation of geospatial information ensures the integration of information systems and the aggregation of distributed network information resources into a single multi-functional ontologically controlled system. The efficiency of the use of information increases due to its timeliness, usefulness, expedient dosage, accessibility (comprehensibility), noise minimization, operational relationship of the source of information

and user, adaptation of the rate of information submission to the rate of its learning, taking into account the individual characteristics of the user, the efficient combination of individual and collective activities, etc. The use of interactive documents of GIS ontology allows significantly expanding the perceptions of such objects of a complex thematic profile of the study, which is the historical and cultural heritage of Kobzar, and interdisciplinary connections between them by supplementing information descriptions of objects based on distributed information resources and searching semantically related information arrays. Such a combination allows us to create a single conceptual information and analytical environment, which is constantly supplemented with the replenishment of geographically distributed users of different directions and has the ability to flexibly expand the functionality through the integration of various information systems.

8 Conclusions

The chapter deals with the approach to structuring the NL text with its key meaningful characteristics on the basis of the primary structure, which is formed during its lexical analysis.

The possibility is determined and the method of structuring natural language texts based on the basis is proposed procedures for recursive text reduction using rules presented in the form of lambda-expressions.

It is shown the possibility of dynamically forming such rules by users without special training on the basis of the original structure of texts already analyzed, which allows to form structured presentation of texts and handle the information contained in the text by automatic and automated systems.

The approach to construct an ontological model of an interactive document is designed to display an expert text structuring results in accordance with entered user request.

The practical implementation of the proposed theoretical approach is a model of ontological GIS-application that is an interactive document.

Such a document is characterized by a natural system of coordinates defined over affine space and due to this, the geospatial information most suitable for displaying naturally.

The models of interactive document and ontological GIS-application provide a high level of representativeness of information available in text documents (in particular, geospatial) for using structured text representation.

Realization of the model of transdisciplinary representation of information as an interactive feature of document, provides the possibility of obtaining operative access to large arrays of thematic information, and in combination with the capabilities of ontological GIS applications - solves the problem of transdisciplinary representation of geospatial information.

References

1. Nicolescu, B.: *Transdisciplinarity - Theory and Practice*, p. 320. Hampton Press, Cresskill (2008)
2. Stryzhak, O.: *Transdisciplinary integration of information resources [text]: author's abstract. dis...* Dr. Tech. Sciences: 05.13.06/Strizhak Alexander Evgenevich; National acad. Sciences of Ukraine, Institute of Telecommunications and Global. Inform Space Kyiv, p. 47 (2014)
3. Kiyaschenko, L., Grebenshikova, E.: *Contemporary Philosophy of Science: Transdisciplinary Aspects*, p. 172., Moscow (2011)
4. Knyazeva, E.: *Transdisciplinary research strategies*. Bull. TGPU **10**, 193–201 (2011)
5. Palagin, A.: *Transdisciplinarity, computer science and development of modern civilization*. Bull. Natl. Acad. Sci. Ukraine **7**, 25–33 (2014)
6. Klein, J.T.: *Transdisciplinarity: A Joint Problem Solving in Science, Technology, and Society: An Effective Way to Managing Complexity*. Birkhäuser (2001)
7. *World Declaration on Higher Education for the Twenty-First Century: Vision and Action*. http://www.unesco.org/education/educprog/wche/declaration_eng.htm. Accessed 02 Dec 2016
8. Novogrudska, R., Globa, L., Koval, O., Senchenko, V.: *Ontology model of intelligent modeling system for marine facilities identification*. In: *Proceedings of International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo)*. IEEE Digital Library (2017). <https://doi.org/10.1109/ukrmico.2017.8095426>
9. Novogrudska, R., Globa, L., Koval, O.: *The simplification method of engineering task sequences used for engineering knowledge portals*. In: *Bulletin of Kharkiv National University named after V. N. Karazin. Mat. modeling. Information Technology. Automated control systems*, № 37, pp. 37–44 (2017)
10. Globa, L., Novogrudska, R., Koval, O., Senchenko, V.: *Ontology for application development*. In: *Ontology in Information Science Ciza Thomas*, IntechOpen. <https://doi.org/10.5772/intechopen.74042>, <https://www.intechopen.com/books/ontology-in-information-science/ontology-for-application-development>. Approved by the Academic Council No. 11. Accessed 27 Nov 2017
11. Globa, L., Novogrudska, R., Koval, A., Senchenko, V.: *Examples of ontology model usage in engineering fields*. In: *Ontology in Information Science CizThomas*, IntechOpen. <https://doi.org/10.5772/intechopen.74369>, <https://www.intechopen.com/books/ontology-in-information-science/examples-of-ontology-model-usage-in-engineering-fields>. Approved by the Academic Council No. 11. Accessed 27 Nov 2017
12. Volvach, I., Globa, L.: *Method of generalized quality index calculation in mobile networks*. In: *Proceedings of 14th International Conference the Experience of Designing and Application of CAD Systems in Microelectronics (CADSM)*. IEEE Digital Library (2017). <https://doi.org/10.1109/cadsm.2017.7916101>
13. Popova, M.: *Ontology of interaction in the environment of the geographic information system*. Thesis of a candidate of technical sciences, Kyiv, p. 250 (2014)
14. Stryzhak, O.E.: *Transdisciplinary integration of information resources: diss. Dr. Tech. sciences, dissertation of the doctor of technical sciences*, Kyiv, p. 470 (2014)
15. Tsvetkov, V.: *Geoinformation systems and technologies*. In: *Finance and Statistics*, Moscow, p. 286 (1998)
16. Andrushchenko, T., Velichko, V., Galchenko, S., Globa, L., Gulyaev, K., Klimova, E., Komova, O., Lisovy, O., Popova, M., Pryhodinuk, V., Stryzhak, O.: *Stus: methods of writing scientific papers on the basis of ontological analysis of texts*. In: *Methodical Manual*. SITIPRINT Ltd., Moscow (2013)

17. Dovgy, S., Popova, M., Prikhodnyuk, V., Stryzhak, O., Yatsukhno, V.: Register of archival documents related to the life, creativity and commemoration of Taras Shevchenko's memory. In: Center for Personal Development "UNICUM", Kyiv, p. 250 (2017)
18. Dovgy, S., Stryzhak, O., Andrushchenko, T., Galchenko, S., Globa, L., Gonchar, V., Kopyyka, V., Kudlyak, V., Lyashuk, K., Popova, M., Pryhodinuk, V., Semeniuk, G., Trofimchuk, O., Polikhun, N., Postova, K.: Ontological study of the life and work of Taras Shevchenko in the environment of the scientific and educational portal KOBZAR.UA. In: Gifted Child Institute, p. 175 (2016)
19. Artemyeva, I., Reshtanenko, N.: Intelligent system based on the multilevel ontology of chemistry. *Softw. Prod. Syst.* **1**, 84–87 (2008)
20. Walkman, Y., Gritsenko, V., Rykhalsky, A.: Model-parametric space: theory and application. In: *Scientific Opinion*, Kyiv, p. 192 (2012)
21. Gavrilova, T., Khoroshevsky, V.: *Knowledge Bases of Intellectual Systems*, p. 384. Peter, St. Petersburg (2001)
22. Gladun, V.: Processes of formation of new knowledge. In: SD "Teacher 6", Sofia, p. 192 (1994)
23. Pospelov, D.: Situational management: theory and practice. In: *Nauka*, Moscow, p. 288 (1986)
24. Ryzhenko, L.: Cognitive engineering. In: *SibADI*, Omsk, p. 172 (2012)
25. Stryzhak, O.: Ontological aspects of transdisciplinary integration of information resources. *Open Inf. Comput. Integr. Technol.* **65**, 211–223 (2014)
26. Guarino, N.: Understanding, building and using ontologies. In: *Human-Computer Studies*, pp. 293–310 (1997)
27. Oletsky, O.: Applying the column "ontology-document" to the problem of intellectual analysis of the behavior of visitors to web resources. In: *Scientific Notes of NaUKMA, Computer Science*, №. 125, pp. 90–92 (2012)
28. Palagin, A., Kryvyi, S., Petrenko, N.: Knowledge-oriented information systems with the processing of natural-language objects: basics of methodology and architectural and structural organization. In: *USiM*, no. 3, pp. 42–55 (2009)
29. Palagin, A., Kryvyi, S., Petrenko, N.: Ontological methods and means of processing subject knowledge: monograph. In: *VNU them. V. Dal*, Lugansk, p. 324 (2012)
30. Palagin, A., Petrenko, N.: System-ontological analysis of the subject domain. In: *USiM*, no. 4, pp. 3–14 (2009)
31. Borovikova, O., Zagorulko, Y.: Organization of knowledge portals on the basis of ontologies. In: *Dialogue 2002: Works of the International Workshop*, Protvino, pp. 76–82 (2002)
32. Velichko, V., Malakhov, K., Semenov, V., Stryzhak, O.: Complex instrumental means of ontology engineering. *Inf. Models Anal.* **3**, 336–361 (2014)
33. Gladun, A., Rogushina, Y.: Ontology in corporate networks. <http://www.management.com.ua/ims/ims115.html>. Accessed 10 June 2017
34. Gorboryukov, V., Stryzhak, O., Franchuk, O.: Use of ontologies in decision-support systems. In: *Mathematical Modeling in Economics: Collection of Scientific Works: S.O. Dovgy (head ed.) [And others]: NAS of Ukraine, Institute of Telecommunications and Global Information Space, Institute of Economics and Forecasting, Institute of Cybernetics named after Glushkov*, Kyiv, pp. 33–39 (2013)
35. Gurzhii, A., Stryzhak, O.: Ontological tools for managing network information resources and their use in educational and scientific activities. In: *Scientific notes of the Small Academy of Sciences of Ukraine: Collection of Scientific Works. Series: Pedagogical Sciences: National Center "Small Academy of Sciences of Ukraine"*, Kyiv, pp. 427–434 (2013)

36. Dobrov, B., Ivanov, V., Lukashevich, N., Solovyev, V.: Ontologies and theses: models, tools, applications: a manual. In: Internet University of Information Technologies; BINOM Laboratory of Knowledge, Moscow, p. 173 (2009)
37. Zagorulko, Y.: Automation of collection of ontological information about Internet resources for the portal of scientific knowledge. In: News of Tomsk Polytechnic University. Management, Computer Engineering and Informatics, pp. 114–119 (2008)
38. Kleshev, A., Salfeeva, E.: Classification of the properties of ontologies. Ontologies and their classification. In: IAPU FEB RAS, Vladivostok, p. 19 (2005)
39. Kotyurova, M.: Stylistics of scientific speech. In: Академия, Moscow, p. 240 (2010)
40. Palagin, A., Kryvyi, S., Velichko, V., Petrenko, N.: To the analysis of natural-language objects. In: “Intelligent Processing” International Book Series “Information Science & Computing”: Supplement to the International Journal “Information Technologies & Knowledge”. ITHEA, Sofia, no. 3. pp. 36–43 (2009)
41. Palagin, A., Petrenko, N.: On the question of system-ontological integration of knowledge of the subject domain. In: Mathematical Machines and Systems, № 3,4, pp. 63–75 (2007)
42. Palagin, A., Petrenko, N.: To designing an ontologically controllable information system with the processing of natural-language objects. Math. Mach. Syst. **2**, 14–23 (2008)
43. Palagin, A., Rippa, S., Sachenko, A.: Conceptualization and problems of ontologies. Artif. Intell. **3**, 374 (2008)
44. Palagin, A., Yakovlev, Y.: System integration of computer equipment. In: UNIVERSUM, Vinnytsia, p. 680 (2005)
45. Palagin, A., Petrenko, N.: System-ontological analysis of the subject domain. In: USiM, no. 4, pp. 3–14 (2009)
46. Popova, M.: Methodology of the formation and use of computer ontologies in the field of ecological education: monograph. In: SITIPRINT, Kyiv, p. 200 (2013)
47. Gruber, T.: A translation approach to portable ontology specifications. In: Knowledge Acquisition, no. 5. pp. 199–220 (1993)
48. Kolmogorov, A., Dragalin, A.: Mathematical logic. In: URSS, Moscow, p. 240 (2005)
49. Stryzhak, O.: Invariant problems of ontological systems. Inf. Technol. Knowl. **8**, 356–360 (2014)
50. Velichko, V., Popova, M., Prikhodnyuk, V., Stryzhak, O.: TODOS - IT-platform for the formation of transdisciplinary informational environments. Armament Mil. Equip. Syst. **1** (49), 10–19 (2017)
51. Velichko, V., Prikhodniuk, V.: Method of automated allocation of relations between terms from natural language texts of technical subjects. In: Knowledge - Dialogue - Solution: Collected Works of the XX International Conference. Q. ITHEA, pp. 27–28 (2014)
52. Velichko, V., Syrota, S., Pryhodnyuk, V.: Instruments of automated distribution of relations from the texts of technical subjects. In: System Analysis and Information Technology: Sb. sciences Works on the materials of the XVI Conference, Kyiv, 26–30 May 2014, p. 348. K.: NTUU “KPI” (2014)
53. Velichko, V., Voloshin, P., Svitla, C.: Automated creation of the thesaurus of the terms of the subject domain for local search engines. In: Knowledge - Dialogue - Solution: International Book Series “Information Science & Computing”. In: ITHEA, Sofia, Bulgaria, no. 15, pp. 24–31 (2009)
54. Velichko, V., Pryhodniuk, V., Stryzhak, A., Markov, K., Ivanova, K., Karastanev, S.: Construction of taxonomy of documents for formation of hierarchical layers in geo-information systems. In: Information Content and Processing, pp. 181–199 (2015)
55. Gladun, V., Velichko, V.: Contemplation of natural-linguistic texts. In: Knowledge - Dialogue - Solution: Proceedings of the XIth International Conference, pp. 344–347 (2005)

56. Korshunova, S.: The role of thesaurus modeling in the organization of terminology “TEXT-TEXT”. In: Herald of the Irkutsk State Linguistic University, no. 1, pp. 116–123 (2009)
57. Malyshevsky, A.: Qualitative models in the theory of complex systems. In: Science. Fizmatlit, Moscow, p. 528 (1998)
58. Prihodniuk, V., Stryzhak, O.: Ontological GIS, as a means of organizing geospatial information. Sci. Technol. Air Forces Armed Forces Ukraine **2**(27), 167–174 (2017)
59. Palagin, A., Velichko, V., Stryzhak, A., Popova: Tools for supporting the analytical activity of the expert in the case study of information resources and sources. Inf. Technol. Knowl. **4**, 329–347 (2010)
60. Popova, M.: A model of the ontological interface of aggregation of information resources and means of GIS. Inf. Technol. Knowl. **7**, 362–370 (2013)
61. Popova, M., Stryzhak, O.: Ontological interface as a means of representing information resources in GIS-environments. In: Scientific notes of the Taurida National University named after VI Vernadsky, Series: Geography, no. 26(65), pp. 127–135 (2013)
62. Mazur, A., Nadutenko, M., Ostapova, I., Petruk, V., Pustovoi, S., Shirokov, V.: Formation of the national terminology system in the field of welding on the basis of virtual lexicographic laboratories. In: Bulletin of the National Academy of Sciences of Ukraine. NAS of Ukraine, Kyiv, pp. 75–83 (2014)
63. Barendregt, X.: Lambda-calculus. His syntax and semantics. In: World, Moscow, p. 606 (1985)
64. Kleshchev, A., Artemyev, I.: Non-enriched systems of logical relations. Sci. Tech. Inf. **7–8**, 18–28 (2000)
65. Lossky, N.: Logic. In: Ripol Classic, p. 167 (1928)
66. Velichko, V., Prihodniuk, V.: Some methods of distinguishing relations between terms in the natural-language text. In: System Analysis and Information Technologies: Collection Sciences Works on the Materials of the XV Conference, Kyiv, 27–31 May 2013, NTUU “KPI”, p. 406 (2013)
67. Pryhodnyuk, V.: Automated formation of electronic layers of geographic information systems on the basis of structured and unstructured information. In: Geoinformation Technologies in Territorial Administration: Collecton of Sciences Works on Materials XIII International Science-Practice Conference, Odessa, 17–18 September 2015, pp. 73–76. Odessa Regional Institute of Public Administration, Odessa (2015)
68. Prihodniuk, V.: Taxonomy of natural-language texts. Inf. Models Anal. **5**, 270–284 (2016)
69. Beklemishev, D.: The course of analytic geometry and linear algebra. In: Higher School, Moscow, p. 320 (1998)
70. Golod, P.: Symmetry and methods of the theory of groups in physics (discrete symmetries). In: Kyiv-Mohyla Academy, Kyiv, p. 215 (2005)
71. Popova, M., Pryhodniuk, V.: Formation of thematic GIS on the basis of semantic-linguistic processing of documents. In: Modern Information Technologies for Environmental Safety Management, Nature Management, Measures in Emergency Situations: Collection of Sciences Works on Materials XIII International. Science-Practice Conference, Kyiv, Puscha-Voditsa, 3–6 October 2016. NAS of Ukraine, Institute of Telecommunications and the Globe. Inform space [and others], Kyiv, Kharkiv, pp. 130–133 (2016)
72. Prihodniuk, V.: Automated selection of information from texts based on the rules presented in λ -expression format. In: Knowledge Management and Competitive Exploration: Collection Sciences of Works on the Materials of the Conference, Kharkiv, 19–21 April 2016. Kharkiv National University of Radio Electronics, pp. 48–49 (2016)

73. Pryhodnyuk, V., Stryzhak, O., Lebid, O.: Ontological representation of the functionality of systems. Ecological safety and nature management: 3 collection of scientific works. In: Voloshkina, O.S., Trofimchuk, O.M. (head eds.) Ministry of Education and Science of Ukraine, Kyiv National University of Construction and Architecture, no. 3–4(22), pp. 5–23 (2016)
74. Pryhodinuk, V., Stryzhak, O.: Multiple characteristics of ontological systems. In: Mathematical Modeling in Economics: 3 Collection of Scientific Works. In: Dovgy, S.O. (head ed.) NAS of Ukraine Institute of Telecommunications and Global Information Space, Institute of Economics and Forecasting, Institute of Cybernetics named after. VM Glushkov, Kyiv, no. 8, pp. 47–61 (2017)



Quality Control for Mobile Communication Management Services in Hybrid Environment

L. Globa^(✉) , M. Skulysh^(✉) , O. Romanov^(✉) ,
and M. Nesterenko^(✉) 

National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”,
Peremohy Avenue 37, Kyiv 03056, Ukraine
lgloba@its.kpi.ua, mskulysh@gmail.com,
a_i_romanov@gmail.com, nikolaiy.nesterenko@gmail.com

Abstract. There is an integration of telecommunication systems and distributed computing environment, resulting in a single hybrid environment for telecommunication services. The hybrid environment has the ability to control the information flow process at each stage and ensures compliance with high quality standards. Providing the quality of service to end-users of communication networks depends on quality control at all stages of the service provision. Today, due to the dynamically changing service structure provided to end users, constantly changing requirements for service quality indicators, with increasing traffic volumes, there is a growing need for well-scalable communication systems that meet the needs of end-users, gaining special significance to service management systems. In the article the features of quality control service servicing of flows of the mobile communication network are investigated with the use of partial virtualization of network functions. An architectural solution for organization of service flows in a hybrid environment is proposed, which includes a telecommunication communication network and cloud computing resources that provide maintenance of virtualized functions involved in the organization of service flows. The solution for improvement of the PCRF system as well as a number of procedures that allow ensuring quality control of servicing streams as well as controlling the computing resources of a hybrid system, which work affects the quality of service of the system service flows, is proposed.

Keywords: NFV · VeCME · VBS · VeEPC · LTE · 5G · TC gibrid-service

1 Introduction

According to [1], hybrid telecommunication service is a service that includes components of cloud and telecommunication services. The work of the telecommunications network is inextricably linked with computer systems. A mobile network consists of a local area network, a radio access network and a provider’s core network. The emergence of cloud computing has expanded the possibilities of servicing telecommunications systems.

Specifications [2] represent the main architectural solutions in which complex hardware solutions are replaced by various ways of virtualizing network partitions. This allows you to flexibly configure network computing resources. To do this, it is necessary to create new methods for managing quality of service, which will take into account the peculiarities of the process in the telecommunications system and in the computing environment for serving hybrid services. This chapter presents the further extension of ongoing research focused on the quality of services provided by mobile operators in the hybrid (cloud and physical) infrastructure. The first approach to the hybrid infrastructure designing has been presented in paper [11].

The range of services and volumes of transmitted traffic are gradually expanding [10, 11, 16–18], which leads to a constant increase in the resources needed to handle the various types of applications that enter the operator's system. Thus, the method of optimal choice of computing resources to serve the variable load of billing systems operator is an urgent task.

Calculation of payment for services requires multi-stage procedures to determine their value, which, as a result, leads to an increase in the load on the tariff system, namely:

- low call service efficiency;
- worsening of both flexibility and efficiency of tariff approaches;
- Failure to guarantee due to quality of service.

A significant number of works and approaches are devoted to customer service quality control. The methods of ensuring quality indicators in the process of information transfer at different stages of service (access to the network, organization of transport information flow) [13–15, 19, 20, 27] are developed, part of the work is devoted to ensuring efficient billing, while the issues of developing PCC rules, in which the policy is set servicing individually for each type of service, depending on the tariff plan [7, 21–23]. Works [24–26] are devoted to the optimization of billing systems connected with the discipline of service aimed at the differentiation of the input stream and the creation of new rules for the sequence of billing subscribers.

However, for telecom operators, the problem of exceeding the connection setup time due to overloading of the network management system is quite acute, because the principle of the distribution of computing resources of technical means, which during periods of peak loads is critical to the quality of service, is not taken into account. Solving this problem may involve recruiting additional leased computing resources. What moments of peak loads will be dynamically involved in processing applications for pricing. This will avoid losses that are due to exceeding the waiting time due to a shortage of computing resources of the system. Exceeding the allowed service life leads to a rejection of the call, in line with economic losses and a reduction in the reputation of the company.

2 Organization of a Heterogeneous Telecommunications Network

According to [3], all computational functions that accompany the transfer process are performed in data centers with cloud infrastructure. Virtualization of the base station will reduce the amount of energy consumed in the dynamic allocation of resources and

load balancing. In addition to virtual base stations, cloud-based radio access networks use their Cloud-RAN resources to create base frequency processing resources that combine various computing resources of a centralized virtual environment. The specification offers virtualization to lose connection to the network for a router located at the border of the provider’s local network. The router performs the classification of flows, routing management and protection of the firewall.

For the organization of virtual base stations and VeCPE, it is necessary that the data center is located close to the base stations and to each LAN output. Thus, the network of the provider is a geographically distributed network of data processing centers with communication channels, which causes the primary information of mobile subscribers to each of them. The network requires conversion at the lowest level, so the signal requires recognition and decoding at higher levels of MAC, RLC, RRC and PDCP. The specification also offers a kernel provider for virtualization.

Therefore, it can be assumed that most of the network processes are performed in data centers, and the network is only a means of delivering informational messages [4, 12]. In the conditions of the distribution of software-managed routers network structure is shown in Fig. 1.

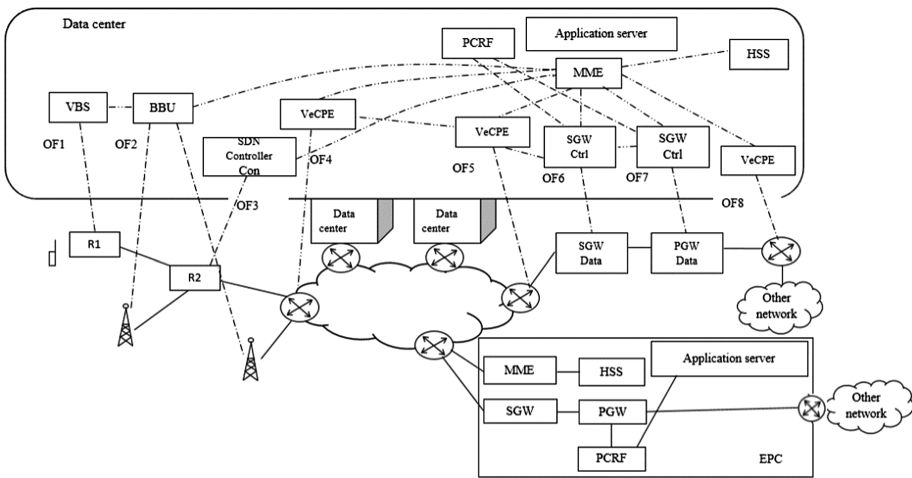


Fig. 1. Provider core network structure using software-controlled routers

Figure 1 shows how the mobile subscriber communicates with the relay R1 that converts the radio signal to the optical one, and then the signal is sent to the repeater R2, controlled by the SDN controller, which is also located in the data center. Upon reaching the data center, the signal is processed by the virtual base station. Further, according to LTE technology, the stream is sent to the core of the operator for further processing. The BBU subsystem is based on the technology of software-configurable networks/network virtualization. This system supports both virtual base stations and the 2G/3G/4G/Pre5G hybrid.

The further direction of the data channels is determined mainly by service. If the thread goes to the ISP's home network, it is immediately sent to the appropriate virtual base station in the data center for service, and then sent to the subscriber through the repeaters R2 and R1. If the thread must be sent outside the operator's local network, it is directed to the marginal virtual router, and then to the external network. This is an example of the next generation network.

Thus, the data center integrates a group of data centers that are connected to a single logical space to serve the functions of the virtualized network through a secure network. The quality of service of end users is influenced by the organization of processes in such a heterogeneous data processing center based on the concept of cloud computing.

According to [28], cloud computing is a network access paradigm that provides a scalable and flexible set of common physical and virtual resources with self-service administration based on demand.

The structure of the described data center, in which the group of functional blocks is shown in Fig. 1 service is shown in Fig. 2. There is a transport network and connected data centers that form a single virtualized space.

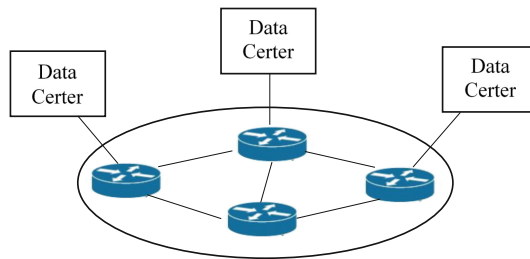


Fig. 2. The structure of the heterogeneous data center

In the Recommendation ITU-T Y.3511 [29], this complex system of data center groups is defined as cloud computing. This is a paradigm for interaction between two or more cloud service providers. Recommendation ITU-T Y.3520 presents a conceptual cloud management architecture and a set of tools written by the services, presented in Fig. 3 [4].

When operating a data provider, the virtual system of the BS, the main subsystems and the virtual router are in the same logical space. Figure 3 shows that at the intermediate software level, the server XXX is represented in each data center that participates in the inter-cloud computing infrastructure. The corresponding programs that activate the functional blocks of the provider are executed at the level of applications and components.

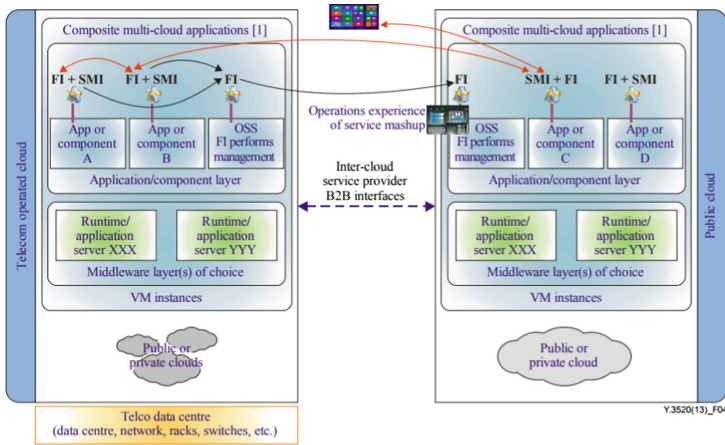


Fig. 3. Architectural vision for multi-cloud, multi-platform cloud management

To ensure the work of the mobile network using virtualization technology, it is necessary to provide a distributed structure of data centers organized in a single virtual space. The structure should include deployed logical elements of the mobile network, process control and flow distribution performed by the orchestra (Fig. 4).

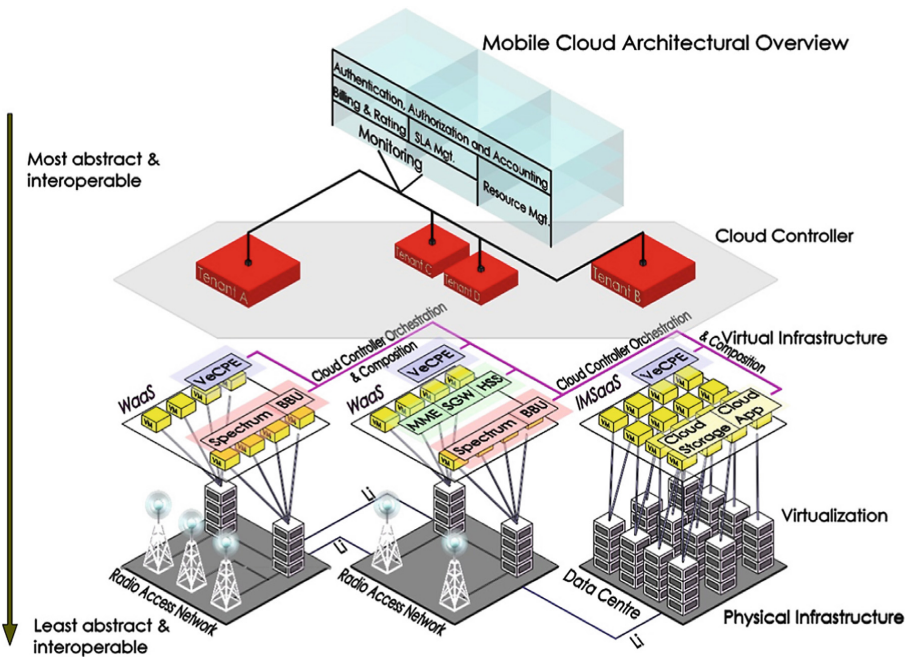


Fig. 4. Organization of service in New Generation Networks

According to the study, the efficiency of the organization of computing processes in functional units affects the efficiency of servicing end users of mobile operators. The data processing center in this architectural solution represents a complex organizational and technical complex of computing and telecommunication systems, ensuring the smooth operation of the NFV infrastructure. The effectiveness of its work depends on the choice of physical data center, which will become part of the structure of the distributed center; location of network functions in the infrastructure; organization of flows between virtualized objects and distribution of resources for their maintenance.

3 The Principle of Flow Service with the Resource Virtualization in Public Telephone Network

Controllers located in the data center, manage all subsystems of mobile communication. The interaction between the controllers of the subsystems for the purpose of control occurs only in the middle of the data center. Functions of the management of the service process, namely: subscriber search, the search for the physical elements involved in the transfer process, and the transfer of instructions to the corresponding physical elements, are transmitted to the cloud.

The subscriber device for interconnecting interacts with the base station controller located in the data center. According to the protocols, the subsystem controllers interact at the data center level, sending the final hardware solutions to the physical equipment to start the data transmission process (Fig. 5).

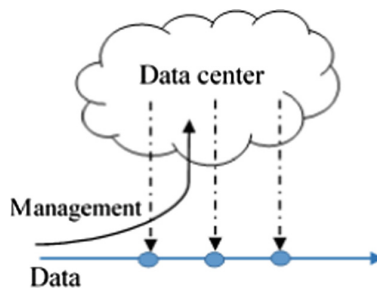


Fig. 5. The principle of flow service with the resource virtualization in public telephone network

There are two principles of virtualization of network resources. The first principle redirects through the cloud resources only control flows. The second principle is to use cloud-based data centers to process both network and information flows. In this paper, the first principle is considered. According to it, virtualization of network functions allows separating the control system of the mobile network nodes from the data

transmission system. In Fig. 6 you can see an example of separating the functions of the core network. The main functions of the core subsystem were analyzed, and the functions associated with the control and data transfer were selected. Data transfer functions are solved in a virtualized environment deployed on the basis of data centers group [5, 6, 8].

- F1 - user filtering and legitimate interception of traffic;
- F2 - IP pool distribution functions for UE and PCEF functions;
- F3 - basic routing and packet traffic interception;
- F4 is a reference point (traffic aggregation point) function for the transfer of service between Node B nodes within the same access network in the base station service area in accordance with a set of rules and instructions;
- F5 - processing functionality of BBERF;
- F6 - the function of detecting traffic;
- F7 - User Data Repository (UDR);
- F8 - program function (AF);
- F9 is an online charging system (OCS).

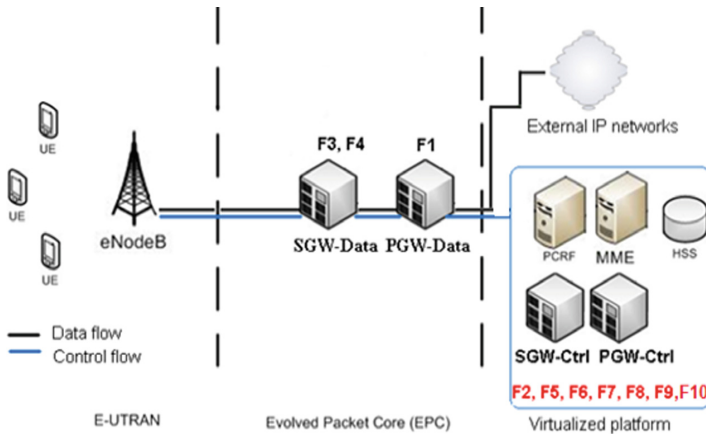


Fig. 6. Distribution of network core functions between physical and virtual devices

Figure 7 shows the processes of interaction of network subsystems with the division of management functions and data transmission during virtualization with the provision of data transfer functions. In fact, each and every arrow in this schema is a maintenance request on this virtual (or physical) node. The number of requests per unit time is the intensity of the load for a given service node.

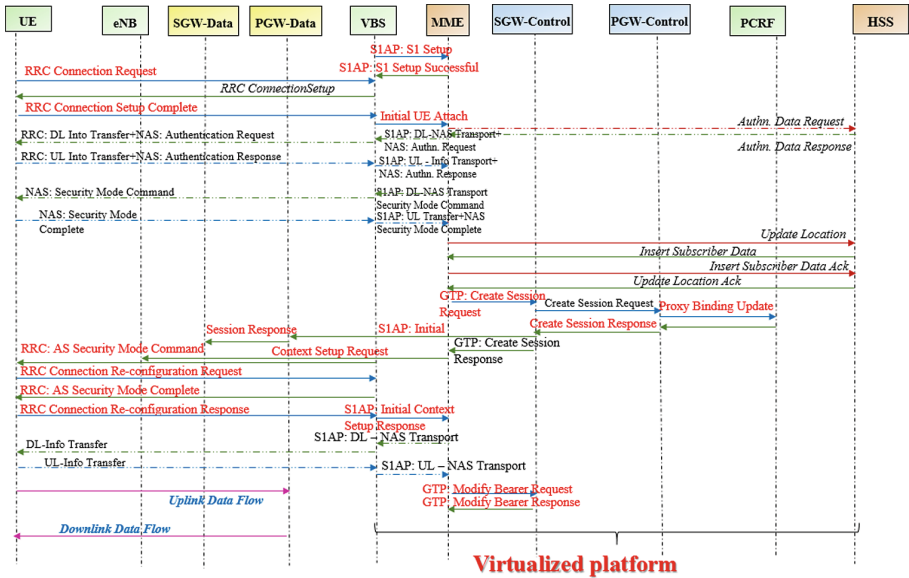


Fig. 7. Procedure of subsystems interaction during subscriber's service

The structure of the network and customer service quality control take place in the nodes. Traditionally, the subsystems of the LTE network perform a set of functions in accordance with standards and specifications. The article proposes to divide subsystem management functions and functions associated with the data transfer process directly into the LTE network. The feature is the expansion of the functionality of subsystems compared with the networks of previous generations. More than half of the functions of the subsystem are connected not with the process of service, but with the management of the communication system. Service quality control occurs in subsystems eNodeB, SGSN, PCRF. Delay management in nodes of a virtualized network, where the intensity of service depends on computing resources, requires modification of the PCRF.

4 Procedure for Ensuring the Quality of Service Provided

The principle of dynamic quality control is as follows: the value of the delay in maintaining the application for connection (disconnect, recovery) is compared with the quality of service of the subscriber. If the metric does not match, then the metrics of quality in virtual nodes and virtual local area networks are compared with the limit values of the corresponding policies stored in the PCRF subsystem. This principle analyzes the following quantitative indicators of the effective operation of the system, such as: the time delay of the service flow request in the virtual node and the probability of loss of queries in the service node. Service node - a virtual machine that performs functions of managing the network node.

After determining the cause of the problem with the performance indicators, appropriate measures are taken. If there is a problem in the time of transmission between service nodes, then it is recommended to reconfigure the system, namely to change the location of virtual nodes in the physical nodes of the heterogeneous structure of the data center. If the problem is detected in a single service node, then it is recommended to increase the number of service resources. If there is a decline in service quality indicators in a group of related interface nodes, for example, which form a single core of the EPC network, then it is recommended to limit the flow of applications sent to service the corresponding kernel. For this purpose it is recommended to calculate the intensity of the load on the group of nodes. Algorithm of procedure of indications on Fig. 8.

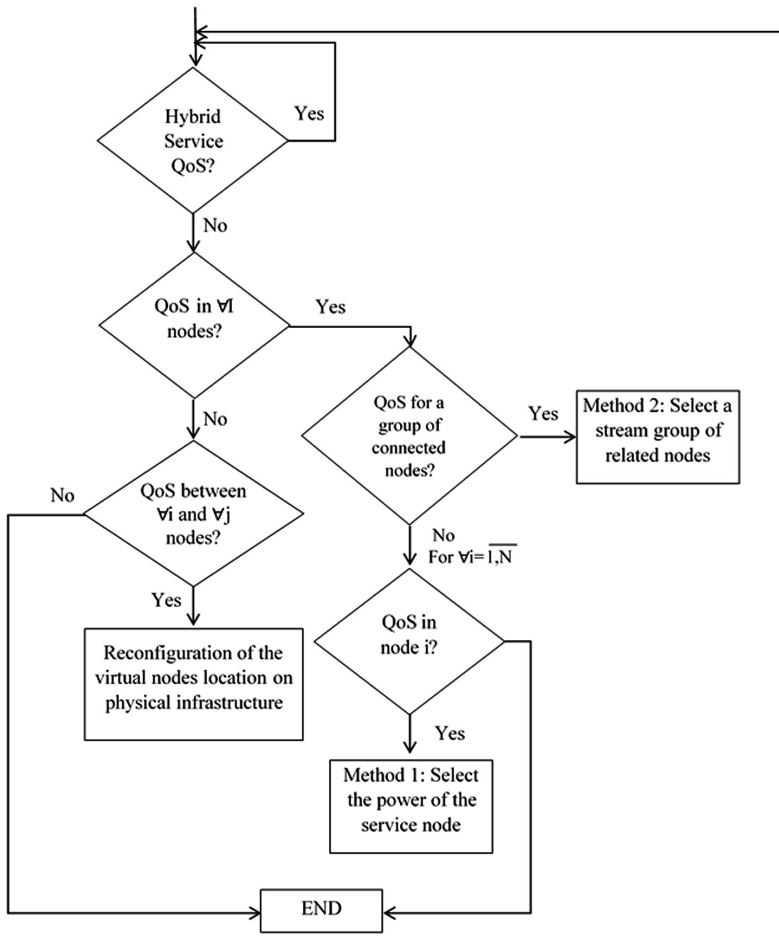


Fig. 8. Procedure for guaranteeing the preset quality of service

To implement the principle of dynamic quality control, a modification of the subsystems of the PCRF system is required. The “Single Policy Storage” subsystem is expanding, and the following policies are added regarding the speed of the flow of quality management services:

1. Permissible latency of service flow in the program on the virtual host.
2. Permissible loss of requests in the virtual node.
3. Permissible time to service queries in groups of virtual nodes that provide this service.
4. Permissible delays in transmission between service nodes.
5. Values of admissible delays of delivery of the directing influence on the network nodes.

An extended subsystem is shown in Fig. 9.

- The “Policy Management” subsystem creates a set of requirements for implementing a set of policies in relation to different flows of management.
- The “Policy Server” subsystem detects a problem of inconsistency between the current quality metrics and the declared subscription policies.
- In the subsystem “Application Server” implemented software modules, in which calculations are performed according to the proposed methods. Output data for the methods is the statistics received from the monitoring system and the policy data provided to the relevant subscribers.
- The subsystem “Subscriber Data Warehouse” is supplemented with information about virtual nodes or a separate database of statistics for the maintenance of a virtual network is created. This database collects information about the flow of service requests; the statistics of the relative dependence of the intensity of service on service resources for each type of request.

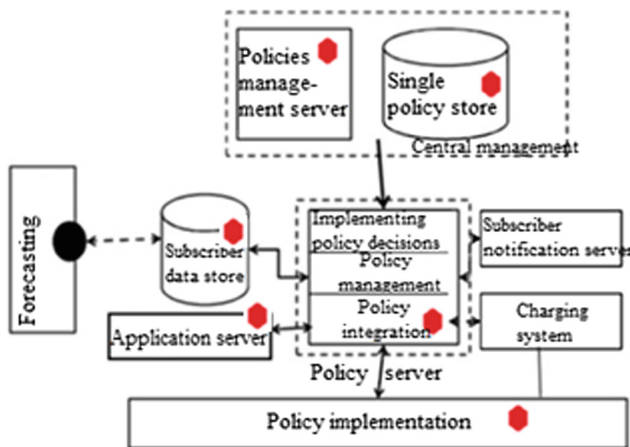


Fig. 9. PCRF subsystem modification

The principle of dynamic quality control requires new procedures: it is necessary to establish the interaction of the mobile communication management system with the virtualized resources management system (Fig. 10).

The quality control of the implementation of management procedures is assessed at the level of the user equipment:

The user logs the delay of the execution of the maintenance procedures, namely the time from the moment of initialization of the connection to the moment of the beginning of data transfer, and passes them to the PCRF subsystem.

PCRF receives this information from the subscriber and analyzes the policy server; In the policy implementation subsystem, he compares the received data with the compliance of the selected subscriber policy that is stored in the Subscriber Data Store.

If the delay values do not match the policy, PCRF asks the Orchestrator subsystem to identify the group of nodes serving the subscriber.

Orchestrator sends the number of nodes that serve the subscriber located in the specified area. The PCRF sends a “Cloud Monitoring” request for information on delay and loss settings in nodes *i* and delay information between service nodes. Cloud monitoring collects information regarding the delay and productivity of loss of hybrid services that are served on nodes in a virtual network. Data on the service node group is transmitted to the PCRF, where the principle of dynamic quality control of hybrid service is implemented. According to the managerial decision, the PCRF subsystem sends inquiries:

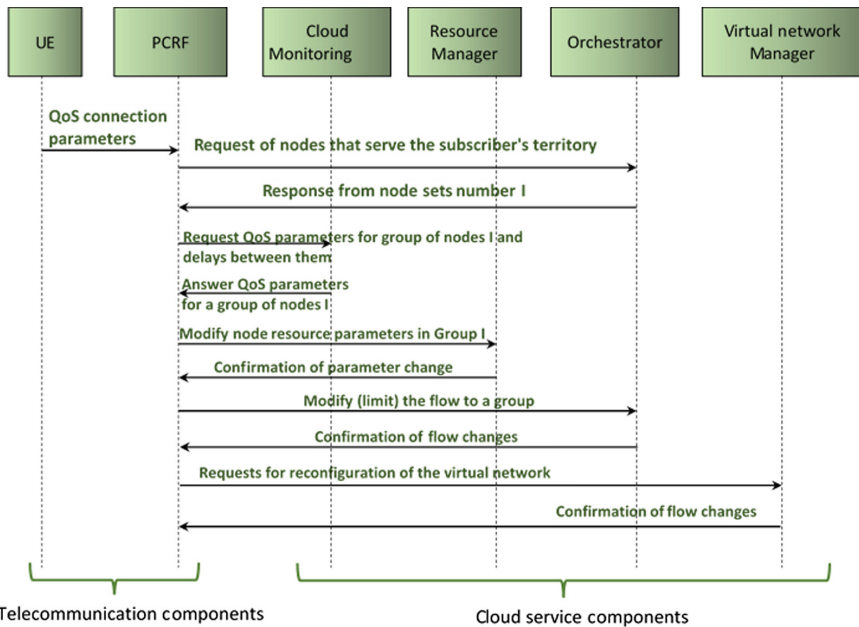


Fig. 10. Interaction of the mobile communication control system and the virtualized environment management system

- to reconfigure the virtual network into “Virtual Network Manager”;
- reconfigure resources to “Resource Manager”;
- Change service streams on Orchestrator streams over a virtual network.

When implementing the principle of dynamic quality control, most subsystems of the PCRF system are involved.

5 Forecasting of the Necessary Virtual Resource in the Cloud to Ensure the Efficient Operation of the Hybrid Telecommunication Environment

On the basis of a large amount monitoring system data, it is necessary to develop service system configurations that will meet the requirements for the quality of the maintenance process. On the basis of configurations and statistics, develop a schedule of resource use and periodically check the sufficiency of resources (Fig. 11).

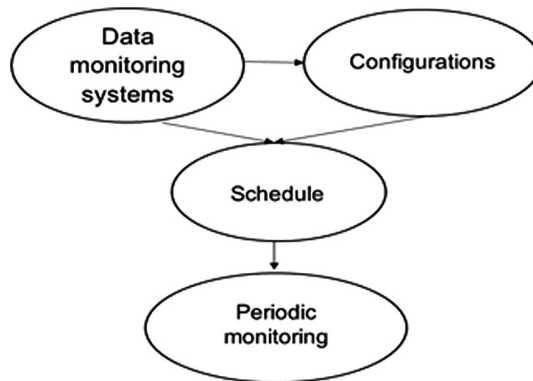


Fig. 11. The principle of forecasting the resource

The main factors influencing the forecast are the unsteadiness of subscriber activity during the day and the possibility of flexible configuration of resources when using the cloud computing environment.

5.1 Method of Constructing a Schedule for Choosing Configurations Based on Long-Term Statistics

The method is based on a large number of statistical data collected over a long period of time (for example, during the year). These data are grouped on different days of the week. It analyzes the long-term load statistics created by the access node group if this method is used to plan the configuration of the service node groups or the load on the service node if the method is used to plan node resource allocation configurations.

The method solves the problem of determining the order of application of service resource configurations, namely the switching points. The basic scientific idea of the method is to use iterative procedures for define the configurations in different time intervals and to check the inequality of Chebyshev to meet the requirements regarding the adequacy of service resources throughout the time line.

Input Data:

- λ_k – maximum service intensity that equals to the k -th configuration;
- $\vec{\lambda}_t$ – vectors of statistical values of the requests quantity, $t \in [0, T]$;
- $\vec{\lambda}_t = (\lambda_t^1, \lambda_t^2, \dots, \lambda_t^{50})$ – this is a set amount values of the requests, which were received at the i -th moment time, during the last 50 weeks.

Output Data:

- (k, t) , where $t \in [0, T]$ – the recommended moments of servicing intensity change,
- k – the recommended configuration of the system, on which should switch to at moment time t .

Algorithm of the Method:

- Step 1.* Split the time axis into time intervals of duration Δt ;
- Step 2.* For each interval $\forall i = \overline{1, \frac{24}{\Delta t}}$ to find the average value of $M\lambda$ and the dispersion σ^2 . For the analysis takes a set of data collected over a long servicing time;
- Step 3.* For each interval to find the minimum configuration k , for which the following inequality is satisfied:

$$M\lambda + 3\sigma^2 \leq \lambda_k \tag{1}$$

- create a temporary set $Kt = \{Kt_i\}$, $Kt_i = (k, i)$, $i = \overline{1, T}$;
- Step 4.* Create a constant set Kt' , if $Kt' = \emptyset$, then $Kt' := Kt$, *Step 5*; else, compare the elements of the sets Kt' i Kt in pairs, if $\forall Kt'_i$ executed $Kt'_i = Kt_i$, then *Step 6*, else $Kt'_i = \max(Kt'_i, Kt_i)$, *Step 5*;
- Step 5.* Change the value Δt , go to *Step 2*;
- Step 6.* Based on Kt'_i to define (k, t) , where $t \in [0, T]$ – the recommended moments of servicing intensity change, k – the recommended configuration of the system, on which should switch to at moment time t .

5.2 Short-Term Load Predicting Method

The short-term load predicting method is improved on the based method ARIMA (Autoregressive integrated moving average) – auto-regression method with sliding mathematical expectation. However, unlike the known method, it is proposed to solve the problem of finding the minimum sliding interval, the use of which will satisfy the requirements, which will minimize the number of floating point operations to perform the prediction, which will ensure the optimal speed of prediction.

The proposed method consists of two stages: the calculation of the prediction interval based on the statistics of the operation of this servicing node and directly periodic forecasting of the load and the control of the sufficiency resources.

Input Data:

- T_p – the time interval for which the prediction is needed;
- λ_i – number of requests per 1 ms, ($i \in 0, \dots, N$), $N = T_{inf}/1$ ms, $\lambda_i \in A$, A – the set of requests number statistics values received during the time T_{inf} (first set, and then adjusted to the 2nd step of the method) before the start of the prediction, $|A| = N$;
- T_{pred} – the prediction period, the time after which the prediction algorithm is started;
- M – the limit of the requests quantity that can be serviced with a given configuration of the servicing device;
- P – likelihood of a prediction error.

Short-term statistics are collected locally at the servicing device and stored no longer than $T_{inf} + T_{pred}$, the interval of sampling is 1 ms.

Output Data:

- $z \in \{0, 1\}$, where $z = 0$ do not change the configuration; $z = 1$ change the configuration 0.

Algorithm of the Method:

Preparatory stage. The tuition of system based on statistical data. The search of the minimum T_{inf} (interval time for data collection)

$$T_{inf} \rightarrow \min, \quad (2)$$

for which the limitation is accomplished:

$$\bar{\lambda}_{T_{pred}} + 3\sigma > M \quad (3)$$

$\bar{\lambda}_{T_{pred}}$ and σ – is calculated according to the main stage.

The limitation is accomplished for $p = P * 100\%$ of statistical samples obtained at different time intervals.

Solution: check values for the sequence established on the principle of $T_{inf}^{k+1} = T_{inf}^k + \Delta$; $T_{inf}^0 = T_{pred}$.

The main stage of dynamic control.

1. To analyse the statistical data λ_i , for the interval T_{inf} time preceding the moment of calculation. Construction by the least square method of coefficients \hat{a} and \hat{b}

$$\lambda = a_i + b. \quad (4)$$

2. To calculate $\bar{\lambda}_{T_{pred}} = \hat{a}T_{pred} + \hat{b}$.
3. If $\bar{\lambda}_{T_{pred}} + 3\sigma \leq M$, then $z = 0$, else $z = 1$.

Planning of HTE resources is based on the analysis of a large number of statistical data. It allows you to calculate servicing configurations and a plan for switching between them during the operation of the control system.

The main stages in planning of the HTE resources are:

- data analysis;
- calculation of configurations;
- calculation of the plan for the use of configurations that ensure the maintenance of the use of resources NFV within the specified limits;
- dynamic control of the sufficiency resources, according to short-term statistics.

Data analysis involves the construction of data flow models for each access point, load statistics model for each servicing node (both virtual and physical), analysis of the flows map over the network (Fig. 12).

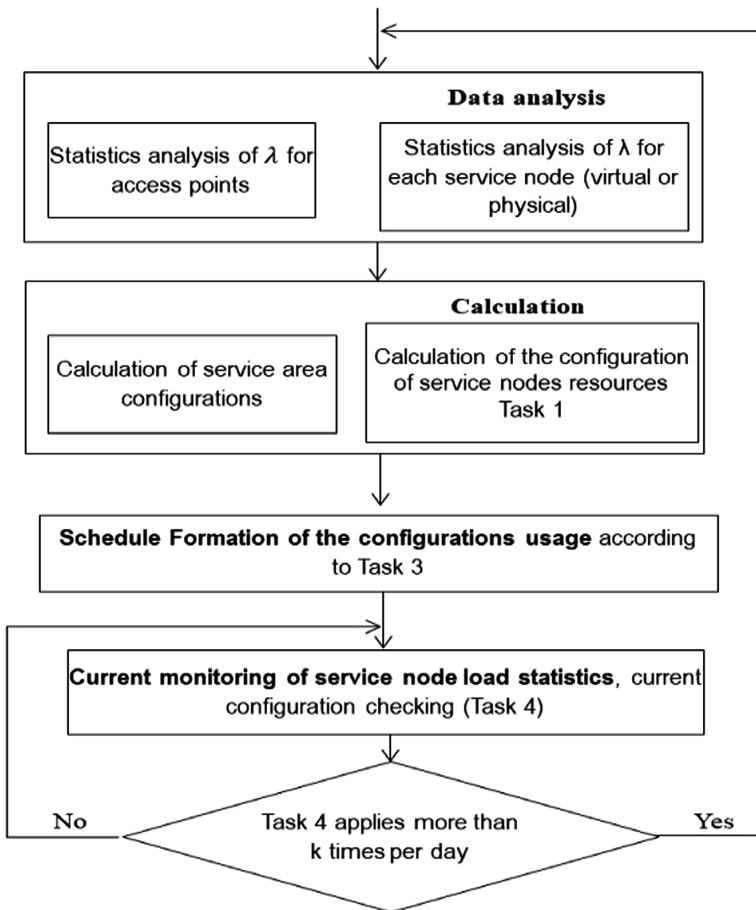


Fig. 12. The algorithm of planning of the HTE resources on the basis of statistical data analysis

Please note that the first paragraph of a section or subsection is not indented. The first paragraphs that follows a table, figure, equation etc. does not have an indent, either. The calculation of the configurations consists of the calculation of the maintenance zones and the calculation of service resources, according to the statistics of loads on the nodes, using [9].

Subsequent paragraphs, however, are indented.

The calculation of service area configurations involves identifying groups of access points that has load statistics, and determining the required configuration of the nodes for servicing the virtualized (or physical) kernel system of the telecom operator.

Then by the method of constructing a schedule on the basis of long-term statistics (p. 5.1), the graph of the use of calculating system configurations is determined.

Further on the method of planning a schedule based on long-term statistics (p. 5.1), the schedule for using the calculating configurations of the system is determined.

To ensure control of overload at service nodes, the sufficiency of resources for the next period of time is monitored on the basis of short-term statistics (p. 5.2).

6 Figure 13 Shows the Basic Procedures for Resource Planning. The Diagram Shows Which Systems Are Involved for Monitoring and Planning Resources

Basic operations are performed in the PCRF system when interoperating between its subsystems namely between “Subscriber data store”, “The application Server”, “Policy Server”, “Order of the policy implementation”, and a heterogeneous environment management system - “Resource Manager” and “Orchestrator”.

Virtual nodes data collection also uses “Virtual entity monitoring” - system that also refers to a heterogeneous environment management system, but is not included in the description of the procedure. It is assumed that the long-term statistics of virtual nodes are stored in the “Subscriber data store”.

Subsequent paragraphs, however, are indented.

Based on the proposed principles, models and methods, a methodology for providing high-quality service of hybrid telecommunication services was proposed, which solves the problem of the lack of tools for monitoring the quality of service in heterogeneous telecommunication environment (HTE) with the virtualization of network functions. In HTE, the quality of service users depends not only on the organization of the telecommunications network, but also on the organization of computing and interaction between the telecommunications and computing environment.

The work systematized all parameters of the HTE and allocated groups:

- group parameters which are changing dynamically, the statistics of which are fixed by monitoring systems cloud service system and telecommunication system;
- groups of sets parameter marginal values specified by standards and specifications and stored on the server by service policies in the PCRF system;
- group of rarely changing parameters;
- group sets of parameters that are calculated by known methods based on the group of parameters described above.

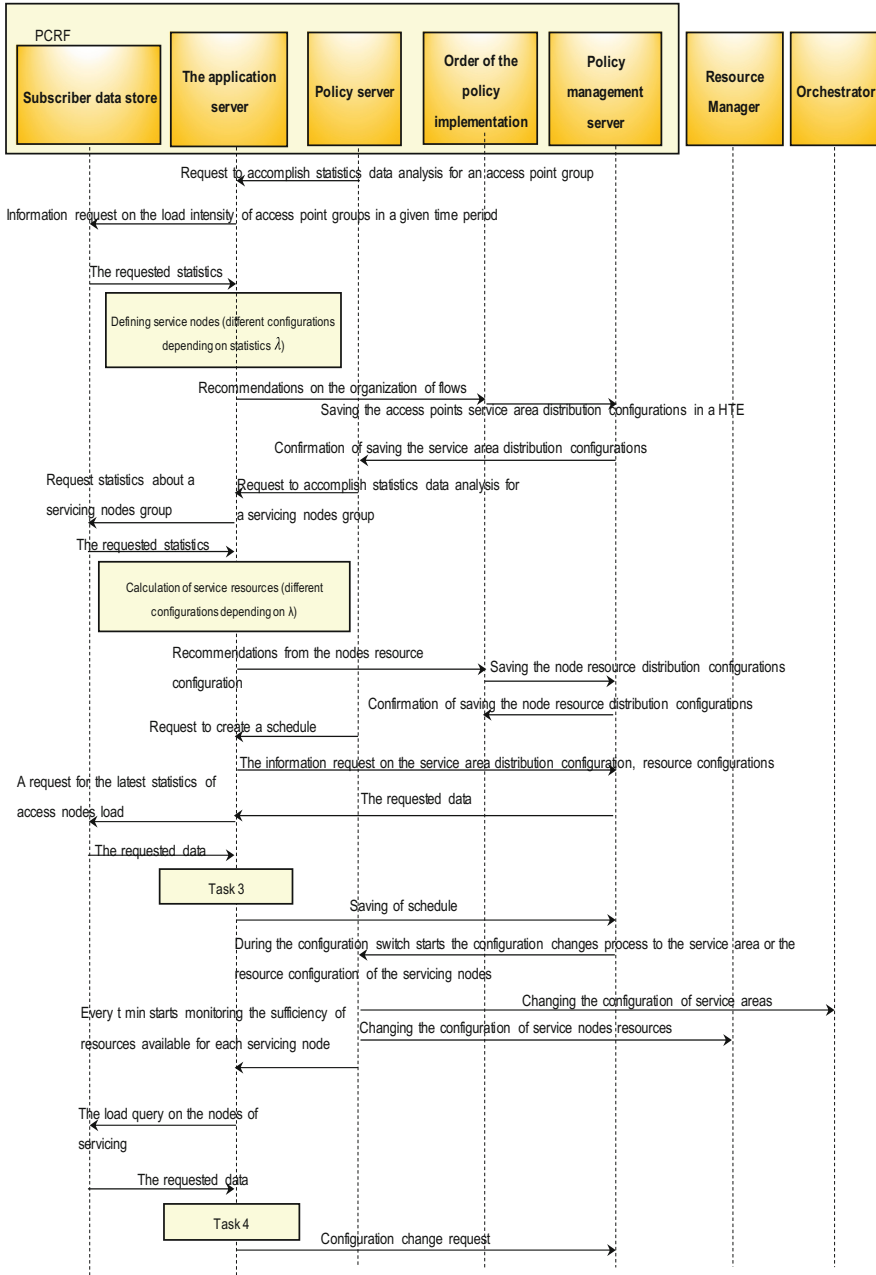


Fig. 13. Procedure for planning the volume of resources of a heterogeneous telecommunication environment

In Fig. 14 the principle of methodology for providing high-quality service of hybrid telecommunication services is given.

Group of dynamically changing parameters:

- $V = \{V\}$ – set of nodes available to service;
- $E = \{e_{ij}\}$ – the matrix of node connectivity at the studied time;
- $A = \{\lambda_{ij}\}$ – the set of the estimated load of the i -th service node;
- $R = \{R_{ij}\}$ – matrix of available resources;
- $N = \{n_{ik}\}$ – set of requests number that simultaneously are serviced at the i -th node;
- $M = \{\mu_i\}$ – set of the service intensities at node.

Group of sets of parameters limit values:

- $D = \{d_{ij}\}$ – set of the delay limit values in the node for various types of services;
- $P = \{P_i\}$ – set of limit values for indicator of packet loss for different types of services;
- $S = \{(S_{mini}, S_{maxi})\}$ – a set of pairs of boundary values of the ratio of optimal use of resources on i -th node.

Paragraphs that follows a table, figure, equation etc. does not have an indent, either.

Group of rarely changed parameters:

- $r = \{k_{ij}\}$ – the possible configurations matrix of node resources, where the i -th – node number, j -th – number of the configuration
- T – set of the time auxiliary parameters used to build the statistical samples and the formation of the schedule.

A group of sets of calculated parameters that are found by known methods:

- M_{max} – set of number of requests that can be served in a node,
- s – set of permissible queues lengths for different types of traffic,
- l – a set of limit values for the number of requests in the queue for which the not applicable method of early system overload,
- $v = \{v_{ij}\}$ – the set of the volume of the j -th type resource that is used to service a single request in the i -th node.

A number of simulation models were presented that implemented the models and methods proposed in the dissertation, on the basis of which the estimation of the quality of service of hybrid telecommunication services was calculated in accordance with the proposed methodology. Service quality assessment was performed on a set of statistical data that was received from the company of the communication operator. The evaluation process proceeded as follows, used modeling the set of kernel nodes of the communication operator in the GPSS system.

Input data of the simulation model:

$dt = 0,1$ ms – interval for time sampling;

$T_{simul} = 864000$ intervals – total simulation time (discrete);

N_{Σ}^i – the number of requests that came in the system for time i -th simulation,

$i = \overline{1,100}$;

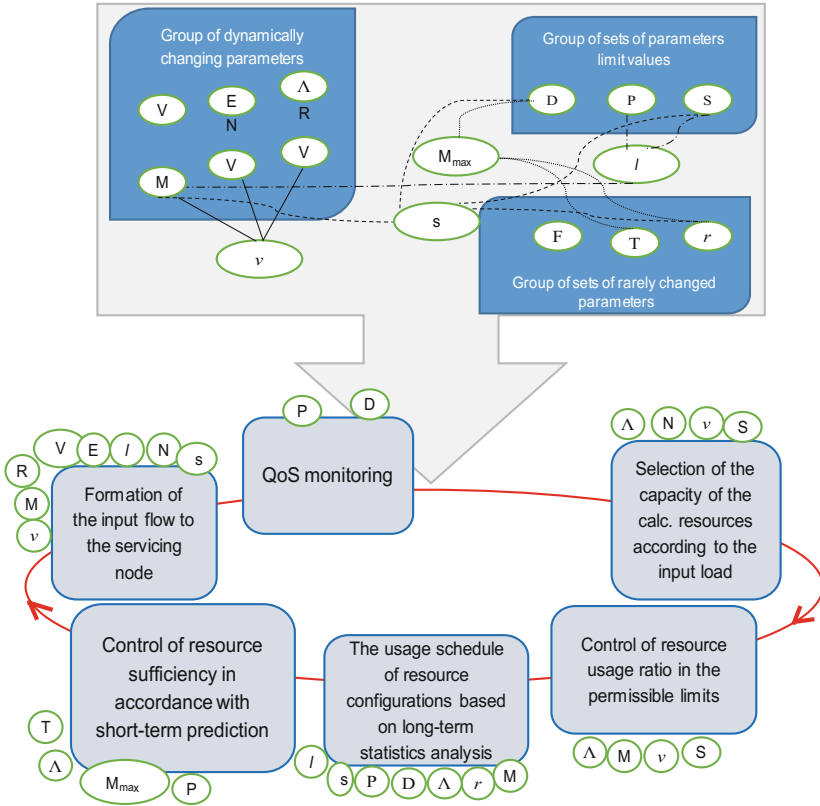


Fig. 14. Principle of methodology for providing high quality service of hybrid telecommunication services

$$M = \begin{cases} 28, & \text{when using additional resources;} \\ 14, & \text{else} \end{cases};$$

$V_j = \{V_{R1j}, V_{R2j}\}$ – available resources of the node j .

For each simulation, the statistics of servicing quality indicators for each j -th node, as well as the system as a whole, was recorded:

t_j^i – average delay time at node j ($j = \overline{1, M}$), during the i -th simulation $i = \overline{1, 100}$;

z_j^i – number of lost queries in node j ($j = \overline{1, M}$), during the i -th simulation $i = \overline{1, 100}$;

t_{all}^i – the average delay time in the system during the i -th simulation $i = \overline{1, 100}$;

z_{all}^i – the number of lost queries in the system during the i -th simulation $i = \overline{1, 100}$.

Also, the monitoring system keep statistic data on the number of resources in each small interval of time:

$R_{1j}^i = \{R_{1j}^1, \dots, R_{1j}^k, \dots, R_{1j}^{T_{MOA}}\}$ – a set data monitoring of resource R1 in node j ($j = \overline{1, M}$), during the i -th simulation $i = \overline{1, 100}$;

$R_{2j}^i = \{R_{2j}^1, \dots, R_{2j}^k, \dots, R_{2j}^{T_{MOA}}\}$ – monitoring of resource R2 in node j ($j = \overline{1, M}$), during the i -th simulation $i = \overline{1, 100}$.

To assess the performance of hybrid services, the calculation of the likelihood of violating the requirements of standards and specifications regarding the service time and the probability of timely service of the service was performed, the corresponding formulas for calculating probabilities were given in Table 1. The simulation results are summarized in Table 2.

Table 1. Quality scores and appropriate quality scores

Quality score	Threshold value	Score	Values of evaluation
t_d – delay time	$P_{ij} = 0,8$ ms	P_{1j} $\forall j = \overline{1, M}$	$P_{1j} = 1 - \left(\sum_{i=1}^{100} k_{ij} \right) / 100$ $k_{ij} = \begin{cases} 1 & t_j^i > P_{ij} \\ 0 & \text{else} \end{cases}$
	$P_{tall} = 8$ ms	P_{1all}	$P_{1all} = 1 - \left(\sum_{i=1}^{100} K_{i1all} \right) / 100$ $K_{i1all} = \begin{cases} 1 & t_{all}^i > P_{tall} \\ 0 & \text{else} \end{cases}$
P – probability of successful service	$P_{zi} = 0,98$	P_{2j} $\forall j = \overline{1, M}$	$P_{2j} = 1 - \left(\sum_{i=1}^{100} K_{iz} \right) / 100$ $K_{iz} = \begin{cases} 1 & \frac{N_{\Sigma}^i - z_j^i}{N_{\Sigma}^i} > P_{zi} \\ 0 & \text{else} \end{cases}$
	$P_{zall} = 0,98$	P_{2all}	$P_{2all} = 1 - \left(\sum_{i=1}^{100} K_{i2all} \right) / 100$ $K_{i2all} = \begin{cases} 1 & \frac{N_{\Sigma}^i - z_{all}^i}{N_{\Sigma}^i} > P_{zall} \\ 0 & \text{else} \end{cases}$
α – resource utilization rate HTE	$a = 0,3$	$P_{3R_{2j}}$ $\forall j = \overline{1, M}$	$P_{3R_{2j}} = \left(\sum_{i=1}^{100} \frac{A_i}{T_{simul}} \right) / 100$
	$b = 0,8$	$P_{4R_{2j}}$ $\forall j = \overline{1, M}$	$P_{4R_{2j}} = \left(\sum_{i=1}^{100} \frac{B_i}{simul} \right) / 100$

In the work of the communication operator, an important indicator of the functioning of the system as a whole is the utilization rate of resources. The practice of the telecommunication company has shown that the resource utilization rate should vary from 30% to 80%. Since, if the resource utilization rate exceeds 80%, unforeseen crashes start, if the resource utilization rate is less than 30%, then hardware is not used effectively and arises surplus of its maintenance costs. Therefore, in the simulation

Table 2. Simulation results

	Standard service	Management based on a methodology	Standard service	Management based on a methodology
	Average delay in servicing the service at the node ($\bar{t}_j = \sum_{i=1}^{100} t_j^i / 100$)		Assessment of timely service in the node ($p_1 = (\sum_{j=1}^M p_{1j}) / M$)	
Max _j	7	9		
Min _j	1	1	0,8	0,805
Average	3,8	4,4		
	Average service servicing delay over the system ($\bar{t}_{all} = \sum_i t_{all}^i / 100$)		Assessment of timely service ($p_2 = p_{1all}$)	
Max	80	90		
Min	20	23	0,82	0,84
Average	55	62		
	Probability of successful servicing in the node ($\bar{z}_j = \sum_i \frac{N_2^i - z_j^i}{N_2^i} / 100$)		Estimation of the probability of successful service in the node ($p_3 = (\sum_{j=1}^M p_{2j}) / M$)	
Max _j	1	1		
Min _j	0,96	0.99	0,95	0.99
Average	0,98	0.999		
	Probability of successful service in the system ($\bar{z}_{all} = \sum_i z_{all}^i / 100$)		Estimation of the probability of successful service in the system ($p_4 = p_{2all}$)	
Max _j	1	1		
Min _j	0,95	0.99	0,94	0.99
Average	0,975	0.999		
	Resource utilization rate HTE α		Probability of using the resource R1 and R2 with a coefficient of utilization of computing resources less than a given threshold value ($\bar{P}_{3R_{1j}}$ i $\bar{P}_{3R_{2j}}$)	
$(\text{Max}_j \bar{R}_{1j}) / V_{R_{2j}}$	0.95	0.9	0.4 i 0,35	0,15 i 0,25
$(\text{Min}_j \bar{R}_{1j}) / V_{R_{2j}}$	0.15	0.25.	Probability of using the resource R1 i R2 with the coefficient of computing resources more than a given threshold value ($\bar{P}_{4R_{1j}}$, $\bar{P}_{4R_{2j}}$)	
$(\text{Max}_j \bar{R}_{2j}) / V_{R_{2j}}$	1	0.85	0,2 i 0,15	0,05 i 0,1
$(\text{Min}_j \bar{R}_{2j}) / V_{R_{2j}}$	0.1	0.2	Estimation of the coefficient of utilization of computing resources used within the norm $p_5 = 1 - \frac{\sum_{g=1}^2 w_g p_{3Rg}}{2} - \frac{\sum_{g=1}^2 w_g p_{4Rg}}{2}$	
Average	0.4	0.2	0.43	0.75

process, the probability that the system resources are used less than a given threshold value a , as well as the probability that the system resources are used more than a given threshold value b

$$R_{2j}^i \supset R_{2j_a}^i = \left\{ R_{2j}^k | R_{2j}^k < a * V_{R2j}; k = \overline{1, T_{simul}} \right\} \quad (5)$$

$$\left| R_{2j_a}^i \right| = A^i \quad (6)$$

$$R_{2j}^i \supset R_{2j_b}^i = \left\{ R_{2j}^k | R_{2j}^k > b * V_{R2j}; k = \overline{1, T_{simul}} \right\} \quad (7)$$

$$\left| R_{2j_b}^i \right| = B^i \quad (8)$$

Table 2 shows that such indicators of quality of service as the average delay in servicing the service at the node and in the system as a whole, the probability of successful servicing of the service in the node and the system as a whole were kept within the limits of acceptable values and acquired a slight improvement. However, the utilization rate of resources for servicing according to the proposed models and methods is kept within the specified limits with an average probability of 32%.

7 Conclusion

The technology of control of the GTS, where the maintenance of hybrid telecommunication services is carried out with the use of software in many cloud-based data-centers, has been proposed, technology has allowed avoiding reducing the quality of service during surges of overload and maintain quality service indicators at a given level, provided that the resource utilization rate is within the specified limits.

The proposed methodology for providing service quality indicators in a hybrid service in a hybrid environment is based on four main approaches: the formation of the load on the service node, the choice of the power of the service node, the determination of the operation of the service nodes and the current control of the system. This allowed ensuring a high-quality functioning of the heterogeneous environment of service of hybrid telecommunication services, namely: to reduce the percentage of requests served in the system of more time allowed by the experts - by 2%, to reduce the percentage of requests lost due to exceeding the allowable service time by 3%, reduce the percentage of time when resources are loaded: less than the permissible value - by 8% or more than the permissible value by 10%.

The article proposes a method for constructing a schedule for attracting resources on the basis of long-term statistics with the provision of service quality indicators for hybrid telecommunication services and the use of service resources within specified limits, which, due to the analysis of long-term statistics, iterative evaluation of statistical characteristics for intervals of different lengths, allows us to formulate a schedule for changing the configuration of resources service for the service node in the GTS, and reduce the resource efficiency by 15%.

The auto-regression method with sliding mathematical expectation is improved, the main difference of which is to determine on the basis of long-term statistics of the interval of analysis of statistics, which provides the reliability of the forecast with a given probability, and minimizes the number of points of analysis, and, respectively, operations for prediction in real time, accuracy of the forecast, compared to a fixed interval of analysis.

The proposed modifications of the components of the PCRf (Policy and Charging Rules Function), new procedures for organizing their interaction with each other and with the virtualization environment control subsystems to ensure the quality control of the service of hybrid telecommunication services, simulation and mathematical modeling of service management processes in the GTS was carried out, which allowed confirming the ability to prevent QoS degradation at the level of the end-user subscriber device and reduce the number of failures due to shortcuts. The response time of the subsystem of control of the telecommunication and cloud portions of the heterogeneous medium.

References

1. ITU-T Recommendation M.3371 of October 2016. Requirements for service management in cloud-aware telecommunication management system. <https://www.itu.int/rec/T-REC-M.3371/en>
2. ETSI GS NFV 001 v.1.1.1 (2013). https://www.etsi.org/deliver/etsi_gs/NFV/001_099/004/01.01.01_60/gs_NFV004v010101p.pdf
3. Skulysh, M., Romonov, O.: The structure of a mobile provider network with network functions virtualization. In: 14-th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering, TCSET 2018, 20–24 February 2018, Conference Proceedings, Lviv–Slavske, pp. 1032–1034 (2018)
4. J. ITU-T Y.3520 Telecommunication standardization sector of ITU (2013). Series Y: Global information infrastructure, internet protocol aspects and next-generation networks
5. Skulysh, M., Klimovych, O.: Approach to virtualization of evolved packet core network functions. In: 2015 13th International Conference on Experience of Designing and Application of CAD Systems in Microelectronics (CADSM), pp. 193–1950. IEEE (2015)
6. Larisa, G., Mariia, S., Andriy, R.: Control strategy of the input stream on the online charging system in peak load moments. In: 2014 24th International Crimean Conference on Microwave & Telecommunication Technology (CriMiCo), pp. 312–313. IEEE (2014)
7. Skulysh, M.: The method of resources involvement scheduling based on the long-term statistics ensuring quality and performance parameters. In: 2017 International Conference on Radio Electronics & Info Communications (UkrMiCo) (2017)
8. Globa, L., Skulysh, M., Sulima, S.: Method for resource allocation of virtualized network functions in hybrid environment. In: 2016 IEEE International Black Sea Conference on Communications and Networking, pp. 1–5 (2016). <https://doi.org/10.1109/blackseacom.2016.7901546>
9. Romanov, O.I., Nesterenko, M.M., Veres, L.A.: IMS: model and calculation method of telecommunication network's capacity. In: Proceedings of the 2017 International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo) 11–15 September 2017, Odessa, Ukraine, pp. 1–4. IEEE Conference Publications (IEEE Xplore Digital Library) (2017). <https://doi.org/10.1109/ukrmico.2017.8095412>

10. Romanov, O.I., Oryschuk, M.V., Hordashnyk, Y.S.: Computing of influence of stimulated Raman scattering in DWDM telecommunication systems. In: 2016 IEEE International Scientific Conference “Radio Electronics and Info Communications”, UkrMiCo 2016 - Conference Proceedings (2016)
11. Skulysh, M.A., Romanov, O.I., Globa, L.S., Husyeva, I.I.: Managing the process of servicing hybrid telecommunications services. Quality control and interaction procedure of service subsystems. In: *Advances in Intelligent Systems and Computing*, vol. 889, pp. 244–256 (2019)
12. Kurdecha, V.V., Zingaeva, N.A.: Optimal reconfigurable base stations (R-BS) architecture and requirements to R-BS. In: 2011 21st International Crimean Conference “Microwave & Telecommunication Technology”, Sevastopol, pp. 465–466 (2011)
13. Ilchenko, M., Uryvsky, L., Moshynska, A.: Developing of telecommunication strategies based on the scenarios of the information community. *Cybern. Syst. Anal.* **53**(6), 905–913 (2017)
14. Moshynska, A., Osypchuk, S., Pieshkin, A., Shmihel, B.: The effect of the features of signalcode constructions forming on indicators of functionality and reliability of communication systems based on the 802.11 N/AC standards. *J. Sci. Europe, Praha, Czech Republic.* № 26, **2**, 38–47 (2018). <http://european-science.org/wp-content/uploads/2018/05/VOL-2-No-26-2018.pdf>. ISSN 3162-2364
15. Chaitanya, T.V.K., Larsson, E.G.: Improving 3GPP-LTE uplink control signaling performance using complex-field coding. *IEEE Trans. Veh. Technol.* **62**(1), 161–171 (2013)
16. Cuevas, A., Moreno, J.I., Einsiedler, H.: IMS service platform: a solution for next-generation network operators to be more than bit pipes. *IEEE Commun. Mag.* 75–81 (2006)
17. Ahson, S.A., Ilyas, M.: *IP Multimedia Subsystem (IMS) Handbook*, p. 250. CRC Press (2009)
18. Ghosh, A., Ratasuk, R., Mondal, B., Mangalvedhe, N., Thomas, T.: LTE-advanced: next-generation wireless broadband technology. *IEEE Wirel. Commun.* **17**(3), 10–22 (2010)
19. Li, X., Bigos, W., Dulas, D., Chen, Y., Toseef, U., Goerg, C., Timm-Giel, A., Klug, A.: Dimensioning of the LTE access network for the transport network delay QoS. In: 2011 IEEE 73rd Vehicular Technology Conference (VTC Spring), pp. 1–7 (2011)
20. Costa-Requena, J.: SDN integration in LTE mobile backhaul networks. In: 2014 International Conference on Information Networking (ICOIN), 10–12 February 2014, pp. 264–269 (2014)
21. Ho, W.-C., Tung, L.-P., Chang, T.-S., Feng, K.-T.: Enhanced component carrier selection and power allocation in LTE-advanced downlink systems. In: 2013 IEEE Wireless Communications and Networking Conference (WCNC), pp. 574–579 (2013)
22. Nuaymi, L., Sato, I., Bouabdallah, A.: Improving radio resource usage with suitable policy and charging control in LTE. In: 2012 6th International Conference on Next Generation Mobile Applications, Services and Technologies (NGMAST), pp. 158–163 (2012)
23. Ouellette, S., Marchand, L., Pierre, S.: A potential evolution of the policy and charging control/QoS architecture for the 3GPP IETF-based evolved packet core. *IEEE Commun. Mag.* **49**, 231–239 (2011)
24. Sou, S.-I., Jeng, J.-Y., Lee, Y.: Signaling overhead of Policy and online Charging Control for bearer sessions in LTE network. In: IEEE 13th International Symposium on Consumer Electronics, ISCE 2009, pp. 593–597 (2009)
25. Malandrino, F., Casetti, C., Chiasserini, C.-F.: LTE offloading: when 3GPP policies are just enough. In: 2014 11th Annual Conference on Wireless On-demand Network Systems and Services (WONS), pp. 1–8 (2014)
26. Gorkemli, B., Tatlicioglu, S., Murat Tekalp, A., Lokman, E.: Dynamic control plane for SDN at scale. *IEEE J. Sel. Areas Commun.* **PP**(99), 1

27. Lira, V., Tavares, E., Oliveira, M., et al.: Virtual network mapping considering energy consumption and availability. <https://doi.org/10.1007/s00607-018-0620-y>
28. ITU-T Recommendation Y.3500: Information technology - Cloud computing - Overview and vocabulary. <https://www.itu.int/rec/T-REC-Y.3500-201408-I/en>
29. ITU-T Recommendation Y.3511: Framework of inter-cloud computing. <https://www.itu.int/rec/T-REC-Y.3511-201403-I/en>



Improved Estimates for the Reliability Indicators of Information and Communication Network Objects with Limited Source Information

D. Mogylevych^(✉)  and I. Kononova^(✉) 

National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Peremoga Avenue 37, Kyiv 03056, Ukraine
mogilevl@email.ua, viti21@ukr.net

Abstract. Analytical relationships are given for quantitative assessment of the lower and upper bounds of reliability indicators for telecommunication objects of two classes: (a) objects with structural redundancy, in which the connection time of backup elements is so small that it can be almost neglected; (b) objects with structural and temporary redundancy, in which connection time is the final (usually random) quantity. The analysis of these relations allowed us to distinguish functionals characterizing the reliability of the objects under consideration, for which exact boundary values were determined. These results made it possible to obtain two-sided assessments of the reliability indicators of the objects of the above classes under conditions when the information on the distribution functions of the connection time of backup elements and the recovery time is limited by knowledge of only two initial moments. Also considered an important case for practice, in which, along with permanent refusals in the equipment of an object, failures may also occur. Analytical relationships have been obtained that allow one to calculate two-sided assessments of the reliability indicators of objects with structural and temporary redundancy, in which equipment failures are taken into account along with refusals.

This chapter concludes with a statement of the results that allow one to find two-way assessments of the reliability indicators of systems with a consistent (in the sense of reliability) connection of objects in which structural or structural and temporary redundancy is provided.

Keywords: Information and communication networks · Refusal · Failure · Redundancy · Readiness index · Reliability index

1 Introduction

The analysis of the processes of information and communication network (ICN) performance shows that the timeliness of the delivery of messages depends on many circumstances, among which the reliability of telecommunication equipment (TE) plays an important role [1–4]. From the point of view of providing and evaluating the reliability of functioning, ICN is characterized by the presence of various types of

redundancy (structural, temporal, informational, functional and derating), introduced into the system and used to ensure normal functioning under the influence of destabilizing external and internal factors. In particular, temporary redundancy is a method for ensuring the reliability of telecommunications objects operating in the transmission of information in conditions of failures, refusals and interference, by assigning and using a backup (redundant) time [5]. Unlike other types of redundancy, there is a time to resign as a reserve.

In a system with an expanding time reserve, the refusal of an object does not lead to a failure of the system's operation, if its recovery time does not exceed the permissible value, which determines the time period used in the system. Also, an assessment of the reliability of functioning at the stages of building and operating systems of the class in question is carried out, as a rule, under a priori uncertainty. This is a typical phenomenon that confronts all stages of the lifecycle, not just ICN.

Telecommunication objects are self-serving functionally-completed ICNs that solve a specific task and ensure the achievement of the common goal of ICN. The implementation of the decomposition principle in the research of reliability of ICN became possible due to the presence of an important characteristic feature of complex systems - the possibility of their partitioning into a finite number of subsystems, and each subsystem, in turn, to a finite number of simpler objects.

Thus, the purpose of the authors of this chapter is to obtain improved estimates of the reliability of ICN objects with limited initial information. The need for such a research is due to the fact that the adoption in many publications of the assumption of complete and reliable source information is not always carried out in practice. More often there is a situation where the distribution functions of determinant random variables are unknown, and only some of their numerical characteristics are known, for example, the first two initial moments. In this case, in general, the exact values of reliability indices are not possible, and the problem consists in finding two-sided estimates (exact lower and upper bounds) of these indices, when unknown functions of the distribution of initial random variables belong to a certain fixed class of distributions. It is for this case that two-sided assessments of reliability of ICN objects are proposed, taking into account failures and persistent refusals.

2 Related Works and Basic Notions

In recent years, a number of scientific publications have appeared, in which a tendency towards the reduction of attention in relation to issues of reliability testing of the functioning of the TE of ICN [2, 6, 7]. Disregard for reliability is often grounded in the fact that TE is modern, and ICN is branched out. This led to the fact that at present there is no single methodology, theoretical positions and effective methods, models, and methods that would allow making informed decisions when modernizing existing and building new ICN.

At the same time, the current stage of the development of TE raises new complex problems, which should take into account the peculiarities of the functioning of the TE of ICN and factors affecting reliability. Among such factors are sustained refusals, failures, a priori statistical uncertainty and different types of redundancy [8, 9].

It should be noted that the overwhelming majority of publications in this subject area contain results that take into account only steady refusals of equipment [10–13]. At the same time, in the last decade there were reports indicating an increase in the number of failures and increasing their impact on the reliability of the TE. In particular, failures in routing protocols is more than 17% of the total number of failures, in power supply systems – up to 16%; in transport networks 84% of refusing is due to failures of optical equipment [14]. A distinctive sign of failure is the operational state of the TE can be avoided without repair, for example, by the operator’s influence on the control bodies. A typical example of a fault is the “freeze” of the TE (router, modem, etc.) or the PC, which is eliminated by a repeated restart program. Damage of a failing character does not affect the operational efficiency of the TE. Each failure of the TE of ICN leads to an interruption of the information exchange for tens of seconds, causing loss of gigabytes of data and, as a consequence, significantly reduces network availability. One of the reasons for this is the widespread use of modern TE computing elements and the use of specialized (applied) software. It should be noted that in most models of reliability TE of ICN, which have been developed to date, only stable refusals of equipment without crashing are taken into account.

By the term “a priori uncertainty” we will understand the existence of incomplete information about the laws of distribution of output random variables that characterize the reliability of individual elements, component parts or ICN as a whole, but also the process of its functioning. The presence of uncertainty of the initial information when solving problems of evaluation and tasks of assessing the reliability indicators caused due to the fact that in many cases it is impossible to obtain a fairly large sample of random variables that characterize failure, repair and the functioning of the system under research, necessary to assess the degree of consistency of the theoretical and statistical distributions. More often, it is possible, for example, to determine the mathematical expectation and dispersion of random variables fairly precisely by the results of tests or according to the data of operation. At the same time, the true distribution function of this random variable can not be determined unambiguously, but it is known that it belongs to the set of all possible theoretical distributions of positive random variables with fixed mathematical expectation and dispersion [15].

In TE, along with structural redundancy, to ensure the reliability of the operation of equipment is widely used time redundancy. In accordance with NSU [16], time redundancy is a redundancy with the use of time reserves. This type of redundancy occurs when the object (TE of ICN) directly in the process of operation provides the ability to spend some time in order to restore the technical characteristics. This time is called backup, and the methods of its use are characterized by time redundancy. This reserve is introduced in order (the algorithm) the use (application) of the object, not in the object, as, for example, in the structural redundancy. Reserve time can be used to detect and locate refusals (failures), restore performance by connecting a structural reserve or repair, reconfiguring the technical structure, changing the network topology, information recovery etc. The peculiarity of this approach to analyzing and ensuring the reliability of complex technical objects is to “weigh” the refusals of importance and a feature of the time spent on eliminating their consequences. This allows us to identify and use internal reserves (in particular, temporal) that are laid down in the objects themselves and their intended use algorithms in order to increase the reliability of the

operation of the objects. Besides, taking into account the time reserve allows you to display true reliability, that is, it allows you to more objectively evaluate the capabilities of systems to function normally under the influence of various destabilizing factors. It is precisely in the presence of a time reserve that in many cases, it can be explained why objects perform their functions more successfully than it produces from the index of equipment refusal [17].

Routers may be examples of the TE of ICN devices that use time redundancy. During performing routing protocols in their algorithms, acceptable (backup) time is provided for eliminating the consequences of TE failures, which determines the time reservation used in ICN as one of the effective ways to increase reliability [18, 19]. In such cases, the disruption of the TE functionality is not necessarily accompanied by a refusal (disruption of functioning), even when the elements are connected sequentially, as there is an opportunity to restore the capacity for the acceptable (backup) time.

The conducted analysis showed [20, 21] that time redundancy as a method of providing the normal functioning of TE in conditions of influence of failures and other destabilizing factors is almost not taken into account in the research of the efficiency of communication equipment, which in some cases leads to the development of mathematical models that inadequately reflect the real level of reliability.

3 Statement of the Problem and Justification of the Choice of the Analytical Method of Solution in the Conditions of a Priori Uncertainty

Let us consider TE, which is a redundant group of elements, which in the following, for brevity, will be called a system. In such a system, routers, switches, or other ICN equipment can be used as elements. Suppose that in general, the system consists of n basic (working) ($n \geq 1$) and m reserve ($m \geq 1$) elements, which for simplicity we will assume to be identical. In the system, together with structural redundancy, temporary redundancy is used, while the structural reserves can have different degrees of load, that is, they can be found in loaded or unloaded states. Let us denote λ and λ_f by that intensity the failures and refusals respectively [22]. We will assume that the kind of functions of the distribution of the time of connection of reserve elements instead of the main rejected $F_c(t) = P\{t_c < t\}$ and the time of repair $F_R(t) = P\{t_R < t\}$ are unknown to us, and only certain of their numerical characteristics are determined, in particular, the first and second initial moments [23].

It is necessary to get the calculated ratios for the quantitative assessment of the basic system reliability indicators. It should be noted that under the above conditions, in the general form, the exact values – of the reliability indices are non-fuzzy, and the problem is to find two-sided estimates (exact lower and upper bounds) of these indices, when unknown distribution functions belong to some fixed class of distributions [24].

To solve this problem, it is expedient to use the fact that in many known settlement ratios for reliability indicators obtained with full source information, there are functionals of a special type (such as linear, fractional-linear etc.) whose values depend on the type functions of distribution of output random variables, with which a priori

information is limited by the value of only the initial moments. If we obtain analytical dependences for the exact lower and upper bounds of these functionalities, then there is a possibility to construct two-way estimates of reliability indicators, which include these functionalities [23].

The above functionality has the form:

$$\left\{ \begin{aligned} I(F) &= \int_0^\infty g(x)dF(x), \\ I(F_1, F_n) &= \int_0^\infty \dots \int_0^\infty g(x_1, \dots, x_n)dF_1(x_1)\dots dF_n(x_n), \\ I(F) &= \frac{I_1(F)}{I_2(F)} = \left(\int_0^\infty g_1(x)dF(x) \right) / \left(\int_0^\infty g_2(x)dF(x) \right), \end{aligned} \right. \quad (1)$$

where $g(x)$ – a given function having finite values at each point x , is limited, can be differentiated, continuous, or piecewise-continuous with finitely many points of discontinuity, and most importantly – depends on the parameters. It is also assumed that for the function $g(x)$ all the points of the discontinuity or angular points, extremum points, bends, etc. are known, that is, the function $g(x)$ is subjected to an analytical research.

We denote by K_1 and K_2 two classes of functions of the distribution of non-negative random variables in accordance with one or two fixed moments

$$s_i = \int_0^\infty x^i dF(x), \quad i = 1, 2, \quad (2)$$

which satisfy the ratio $0 < s_1^2 < s_2 < \infty$. The subset of distribution functions from classes K_1 and K_2 satisfying an additional condition $F(Q + 0) = 1$ (where Q is a fixed smallest number satisfying this condition), we denote by $K_i^Q, i = 1, 2$. In classes K_i^Q there are inequalities $0 < s_1^2 < s_2 < s_1 Q$.

With specified restrictions on the function $g(x)$ and moments s_1 and s_2 and functional (1) and their exact upper and lower boundaries exist. If the function $g(x)$ is continuous, then integrals (1) exist as Stieltjes integrals. If both functions $g(x)$ and $F(x)$ are discontinuous at the same point, then the Stieltjes integral, as is known, does not exist. In this case, the integrals (1) exist as the Lebesgue-Stieltjes integrals [25].

You need to consider tasks such as:

$$I(F) \rightarrow \sup, F \in K,$$

or

$$I(F) \rightarrow \inf, F \in K, \quad (3)$$

which means to find supremum or infimum of the corresponding functional when F belonging to the corresponding class K .

Currently, two basic approaches to solving the above tasks have been determined. The basis of the first approach is a numerical method based on the ideas of linear stochastic programming [26, 27].

The second approach involves an analytic solution to the problem of finding the extreme distribution functions for the calculation of the lower and upper bounds of the functional (1), in which the subintegral functions depend on the parameters [28–30].

Despite the fact that to date, the general theory and methods for solving such problems have been developed quite fully, the finding of extremums and boundary distributions for specific functionalities by analytical methods, in the figurative expression of Academician I.N. Kovalenko, is more an art [31]. Therefore, there is a need for further development in the application plan to determine the two-sided estimates (exact limit values) of the functionality that characterizes the reliability of ICN objects.

4 Improved Method of Constructing Two-Sided Assessments of Reliability of TE Taking into Account Refusals and Failures

The essence of the proposed approach consists in the allocation of typical functional, through which the main indicators of reliability of TE objects with time, structural and derating redundancy are expressed, obtained with full initial information [15, 23, 32], and the construction of the exact lower and upper bounds of these functionals with known initial moments of distribution of individual output random variables, followed by obtaining two-sided estimates (lower and upper bounds) of reliability indicators by substituting them for the boundary values of the corresponding functionals [33].

Thus, the implementation of this approach includes the implementation of the following steps: the choice of typical functional, characterizing the reliability of backup systems; obtaining for these functionals the exact lower and upper bounds; construction of two-sided assessments (lower and upper bounds) of reliability of reserved systems [34].

The analysis showed that the most difficult task is the implementation of the first two stages. Let's consider the solution of this problem in more detail.

The analysis of the known results of the research of the reliability of reserved systems [34] has allowed distinguishing three types of functionals, which are included in reliability indicators:

$$I_1(F_c) = F_c(t_a) = \int_0^{t_a} g(x, t_a) dF_c(x), \quad (4)$$

where

$$g(x, t_a) = \begin{cases} 1, & 0 \leq x < t_a, \\ 0, & x \geq t_a; \end{cases}$$

$$I_2(F_a) = \int_0^{t_a} [1 - F_c(x)] dx; \tag{5}$$

$$I_3(F_c) = \frac{\int_0^{t_a} x dF_c(x)}{1 - F_c(t_a)} - t_a. \tag{6}$$

The function of the time division of connection of elements of the structural reserve $F_c(x)$ included in these functions belongs to the set of distribution functions K_2 that satisfy the relation (2) and the constraints $0 < s_1^2 < s_2 < \infty$.

It is necessary to find the exact lower $\inf_{F_c \in K_2} I_i(F_c)$ and upper $\sup_{F_c \in K_2} I_i(F_c)$ bounds of the functionals $I_i(F_c)$, $i = 1, 2, 3$. Notation $F_c \in K_2$ means that the unknown distribution function $F_c(x)$ belongs to the set of functions of the distribution $K_2(s_1, s_2)$ of positive random variables with fixed initial moments s_1 and s_2 .

The derivation of the calculation formulas for the evaluation of the limit values of the functional $I_i(F_c)$, $i = 1, 2, 3$, is given in [33].

The lower and upper bounds of the functional $I_1(F_c)$ (4) are given in Table 1.

The lower and upper bounds of the functional $I_2(F_c)$ (5) have the form:

Table 1. Lower and upper bounds of the functional $I_1(F_c)$

Change the setting t_a	$\inf_{F_c \in K_2} I_1(F_c)$	Growth points of the boundary distribution function $F_0(x)$	Boundary polynomials $U_0(x)$	$\sup_{F_c \in K_2} I_1(F_c)$	Growth points of the boundary dis-tribution function $F_0(x)$	Boundary polynomials $U_0(x)$
$0 < t_a < s_1$	0	$x_1 = t_a,$ $x_2 = n,$ $n \rightarrow \infty$	$U_0(x) = 0$	$\frac{s_2 - s_1^2}{s_2 - 2s_1 t_a + t_a^2}$	$x_1 = t_a,$ $x_2 = \frac{s_2 - s_1 t_a}{s_1 - t_a}$	$U_0(x) = \frac{(x - x_2)^2}{(t_a - x_2)^2}$
$s_1 \leq t_a < \frac{s_2}{s_1}$	$\frac{t_a - s_1}{t_a}$	$x_1 = 0,$ $x_2 = t_a$	$U_0(x) = 1 - \frac{x}{t_a}$	1	$x_1 = 0,$ $x_2 = t_a$	$U_0(x) = 1$
$t_a \geq \frac{s_2}{s_1}$	$\frac{(t_a - s_1)^2}{s_2 - 2s_1 t_a + t_a^2}$	$x_1 = \frac{s_1 t_a - s_2}{t_a - s_1},$ $x_2 = t_a$	$U_0(x) = 1 - \frac{(x - x_1)^2}{(t_a - x_1)^2}$	1	$x_1 = \frac{s_1 t_a - s_2}{t_a - s_1},$ $x_2 = t_a$	$U_0(x) = 1$

$$\inf_{F_c \in K_2} I_2(F_c) = \begin{cases} t_a (s_1^2 / s_2), & t_a < s_2 / 2s_1, \\ 0, S \left(s_1 + t_a - \sqrt{t_a^2 - 2t_a s_1 + s_2} \right), & t_a \geq s_2 / 2s_1; \end{cases} \tag{7}$$

$$\sup_{F_c \in K_2} I_2(F_c) = \sup_{F_c \in K_2} \int_0^{t_a} [1 - F_c(x)] dx = \begin{cases} t_a, & t_a < s_1, \\ s_1, & t_a \geq s_1. \end{cases} \tag{8}$$

The lower and upper bounds of the functional $I_3(F_c)$, $F_c \in K_2$, $F_c \in K_2^Q$ (6) are given in Tables 2 and 3.

Table 2. Exact lower estimates of the functional $I_3(F_c)$, $F_c \in K_2$, $F_c \in K_2^Q$

Parameter change areas	$0 < t_a < s_1$	$s_1 < t_a < s_2/s_1$		$t_a \geq s_2/s_1$
	$Q \leq \infty$	$Q = \infty$	$Q < \infty$	$Q \leq \infty$
Growth points of extreme distribution functions	$x_1 = t_a + 0$ $x_2 = B(t_a)$	$x_1 = 0$ $x_2 = t_a + 0$	$x_1 = 0$ $x_2 = t_a + 0$ $x_3 = Q$	$x_1 = t_a + 0$ $x_2 = B(t_a)$
$\inf I_3(F_c)$	$s_1 - t_a$	0	$\frac{t_a(s_2 - s_1 t_a)}{(Q + t_a)s_1 - s_2}$	$s_1 - t_a$

Table 3. Exact upper bounds of the functional $I_3(F_c)$, $F_c \in K_2$, $F_c \in K_2^Q$

Parameter change areas t_a and Q	$Q \leq \infty$ $0 < t_a < s_1$	$Q < \infty$ $s_1 < t_a < s_2/s_1$
	Growth points of extreme distribution functions	$x_1 = t_a - 0$ $x_2 = B(t_a)$
$\sup I_3(F_c)$	$B(t_a) - t_a = \frac{s_2 - 2s_1 t_a + t_a^2}{s_1 - t_a}$	$Q - t_a$

Note. In Table 3 accepted designations: $B(t_a) = \frac{s_2 - 2s_1 t_a}{s_1 - t_a}$,

$$B(Q) = \frac{Qs_1 - s_2}{Q - s_1}.$$

Thus, a set of linear functionals (4)–(6) is defined that characterize the reliability of TE objects with different types of redundancy, and the calculated relations are obtained for finding the lower and upper bounds of these functionals (Tables 1, 2 and 3 and formulae (7), (8)) [34]. In the following sections, the resulting boundary values of these functionals will be used when constructing two-sided assessments of the reliability of the TE objects with time, structural and derating redundancy.

5 Two-Sided Estimates of the Reliability of TE Taking into Account Refusals and Failures when Using Different Types of Redundancy

Obtained in the previous section, the boundary values of the functionals that characterize the reliability of TE objects with a replenished time reserve, allow us to determine two-sided assessments of the reliability of this class of systems. To do this, it is

necessary to implement the third stage of the proposed method of solving the problem: to set the boundary values of the functional in the calculated relations for reliability indicators, which are obtained with full source information.

5.1 Models of Reliability of TE with Time, Structural and Derating Redundancy

It should be noted that in the case under consideration, the elements of the structural redundancy are in an unloaded state. The probability of refusal (failure of functioning) of the system under consideration when connecting a backup element to the main one, which is rejected, is expressed by the formula:

$$q = 1 - F_c(t_a) = 1 - I_1(F_c), \tag{9}$$

where $I_1(F_c)$ is the functional for which the exact boundary values are obtained (Table 1).

Using (9) and the formulas of Table 1, it is easy to get the lower q_* and upper q^* bound of the probability q :

$$q_* = 1 - \sup_{F_c \in K_2} I_1(F_c),$$

$$q^* = 1 - \inf_{F_c \in K_2} I_1(F_c).$$

The formulas for q_* and q^* for various areas of parameter t_a change are given in Table 4.

Using the limit values of probability q (Table 4), we obtain two-sided estimates of system reliability: average failure rate $T_0(t_a)$ and probability of failure-free operation $P(t, t_a)$:

$$\min_{F_c \in K_2} T_0(t_a) = \frac{1}{n\lambda} \left[\frac{x(n\rho)^m}{m!} + q^*(1 + k_f) \right]^{-1}, \tag{10}$$

$$\max_{F_c \in K_2} T_0(t_a) = \frac{1}{n\lambda} \left[\frac{x(n\rho)^m}{m!} + q_*(1 + k_f) \right]^{-1}; \tag{11}$$

$$\max_{F_c \in K_2} T_0(t_a) = \frac{1}{n\lambda} \left[\frac{x(n\rho)^m}{m!} + q_*(1 + k_f) \right]^{-1}; \tag{11}$$

Table 4. Limit values q_* and q^* probability of refusal q

Parameter change areas t_a	Lower boundary q_*	Upper boundary q^*
$0 < t_a < s_1$	$\frac{(s_1 - t_a)^2}{s_2 - 2s_1 t_a + t_a^2}$	1
$s_1 \leq t_a < \frac{s_2}{s_1}$	0	$\frac{s_1}{t_a}$
$t_a \geq \frac{s_2}{s_1}$	0	$\frac{s_2 - s_1^2}{s_2 - 2s_1 t_a + t_a^2}$

$$\min_{F_c \in K_2} P(t, t_a) = \exp \left[-\frac{t}{\min_{F_c \in K_2} T_0(t_a)} \right] = \exp \left\{ -n\lambda t \left[\frac{x(n\rho)^m}{m!} + q^*(1 + k_f) \right] \right\}, \quad (12)$$

$$\max_{F_c \in K_2} P(t, t_a) = \exp \left[-\frac{t}{\max_{F_c \in K_2} T_0(t_a)} \right] = \exp \left\{ -n\lambda t \left[\frac{x(n\rho)^m}{m!} + q_*(1 + k_f) \right] \right\}, \quad (13)$$

where q_* and q^* – respectively, the lower and upper limits of the probability of refusal of the system (see Table 4); $\rho = \lambda \bar{t}_R$; $k_f = \frac{\lambda_f}{\lambda}$, $x = \begin{cases} \frac{\beta_m}{\beta_1^m}, & l = 1, \\ 1, & l \geq m. \end{cases}$

Complex indicators of system reliability include a functional $I_3(F_c)$. Significantly through $\underline{T}_c^*(t_a)$ and $\bar{T}_c^*(t_a)$ the lower and upper bounds of this functional and we will assume that the random variable t_c with the distribution function $F_c(t)$ can not take values greater than Q , that is, $F_c(Q + 0) = 1$, $F_c(Q - 0) < 1$ and $0 < s_1^2 < s_2 < s_1 Q$.

Then

$$\underline{T}_c^*(t_a) = \inf_{F_c \in K_2^Q} I_3(F_c) = \begin{cases} s_1 - t_a, & 0 < t_a < s_1, \\ \frac{t_a(s_2 - s_1 t_a)}{(Q + t_a)s_1 - s_2}, & s_1 < t_a < \frac{s_2}{s_1}, \\ 0, & t_a \geq \frac{s_2}{s_1}, \end{cases} \quad (14)$$

$$\bar{T}_c^*(t_a) = \sup_{F_c \in K_2^Q} I_3(F_c) = \begin{cases} \frac{s_2 - 2s_1 t_a + t_a^2}{s_1 - t_a}, & 0 < t_a < s_1, \\ Q - t_a, & s_1 < t_a < \frac{s_2}{s_1}. \end{cases} \quad (15)$$

After the replacement of the limit values $\underline{T}_c^*(t_a)$ and $\bar{T}_c^*(t_a)$ determine the two-sided estimates of the average time of system recovery [34]:

$$\min_{F_c \in K_2} T_R(t_a) = \frac{1}{\Lambda_1 + \Lambda_2} [\Lambda_1 T_R^* + \Lambda_2 \underline{T}_c^*(t_a)], \quad (16)$$

where

$$\Lambda_1 = n\lambda \left[\frac{x(n\rho)^m}{m!} \right], \quad \Lambda_2 = q_* n\lambda (1 + k_f);$$

$$\max_{F_c \in K_2} T_R(t_a) = \frac{1}{\Lambda_1 + \Lambda_3} [\Lambda_1 T_R^* + \Lambda_3 \bar{T}_c^*(t_a)], \quad (17)$$

where

$$\Lambda_3 = q^* n\lambda (1 + k_f).$$

In (16) and (17), $T_R^* = \begin{cases} \frac{\beta_{m+1}}{(m+1)\beta_m}, & l = 1, \\ \frac{\beta_1}{m+1}, & l = m + 1; \end{cases}$, q_* and q^* – respectively, the lower and upper limits of the probability of refusal $q = 1 - F_c(t_a)$ when connecting the backup element.

Using (16) and (17), it is easy to obtain two-sided assessments of complex system reliability indices – readiness coefficient and readiness coefficient [5, 34]:

$$\min_{F_c \in K_2} K_h(t_a) = \left[1 + \frac{\max_{F_c \in K_2} T_R(t_a)}{\min_{F_c \in K_2} T_0(t_a)} \right]^{-1}, \tag{18}$$

$$\max_{F_c \in K_2} K_h(t_a) = \left[1 + \frac{\min_{F_c \in K_2} T_R(t_a)}{\max_{F_c \in K_2} T_0(t_a)} \right]^{-1}, \tag{19}$$

$$\min_{F_c \in K_2} P_g(t, t_a) = \min_{F_c \in K_2} K_h(t_a) \exp \left[-\frac{t}{\min_{F_c \in K_2} T_0(t_a)} \right], \tag{20}$$

$$\max_{F_c \in K_2} P_g(t, t_a) = \max_{F_c \in K_2} K_h(t_a) \exp \left[-\frac{t}{\max_{F_c \in K_2} T_0(t_a)} \right]. \tag{21}$$

Two-sided estimates of system reliability for a particular case $n = m = 1$ (duplicate system) are given in Table 5.

Table 5. Duplicate system with time and derating redundancy

Indicator	Limit values	Estimated formulas
$T_0(t_a)$	$\min_{F_c \in K_2} T_0(t_a)$	$(1/\lambda)[\rho + q^*(1 + k_f)]^{-1}$
	$\max_{F_c \in K_2} T_0(t_a)$	$(1/\lambda)[\rho + q_*(1 + k_f)]^{-1}$
$P(t, t_a)$	$\min_{F_c \in K_2} P(t, t_a)$	$\exp[-\lambda t(\rho + q^*(1 + k_f))]$
	$\max_{F_c \in K_2} P(t, t_a)$	$\exp[-\lambda t(\rho + q_*(1 + k_f))]$
$T_R(t_a)$	$\min_{F_c \in K_2} T_R(t_a)$	$(\rho T_R^* + q^*(1 + k_f)\underline{T}_c^*(t_a))/(\rho + q^*(1 + k_f))$
	$\max_{F_c \in K_2} T_R(t_a)$	$(\rho T_R^* + q_*(1 + k_f)\bar{T}_c^*(t_a))/(\rho + q_*(1 + k_f))$
$K_h(t_a)$	$\min_{F_c \in K_2} K_h(t_a)$	$[1 + \lambda(\rho T_R^* + q_*(1 + k_f)\bar{T}_c^*(t_a))]^{-1}$
	$\max_{F_c \in K_2} K_h(t_a)$	$[1 + \lambda(\rho T_R^* + q^*(1 + k_f)\underline{T}_c^*(t_a))]^{-1}$
$P_g(t, t_a)$	$\min_{F_c \in K_2} P_g(t, t_a)$	$\min_{F_c \in K_2} K_h(t_a) \exp[-\lambda t(\rho + q^*(1 + k_f))]$
	$\max_{F_c \in K_2} P_g(t, t_a)$	$\max_{F_c \in K_2} K_h(t_a) \exp[-\lambda t(\rho + q^*(1 + k_f))]$

Note: $T_R^* = \beta_2/2\beta_1$ when $l = 1$; $T_R^* = \beta_1/2$ when $l = 2$; $\underline{T}_c^*(t_a)$ – formula (14); $\bar{T}_c^*(t_a)$ – (15); q_* and q^* – Tables 3, 4; $\beta_1 = \bar{t}_R$; $\rho = \lambda \bar{t}_R$; $k_f = \lambda_f/\lambda$.

5.2 Reliability Models of TE with Time and Structural Redundancy

In this system, the elements of the structural redundancy are in a loaded state. Here are the calculated ratios for reliability indicators. Using the limit values of the probability q of refusal of the system (failure of functioning) when connecting elements of the structural reserve (Table 4), we take into account two-sided assessments of the reliability of the system [17]:

$$\min_{F_c \in K_2} T_0(t_a) = \frac{1}{n\lambda} \left[\frac{x\rho^m}{m!} \prod_{i=1}^m (n+i) + q^*(1+k_f) \right]^{-1}, \quad (22)$$

$$\max_{F_c \in K_2} T_0(t_a) = \frac{1}{n\lambda} \left[\frac{x\rho^m}{m!} \prod_{i=1}^m (n+i) + q_*(1+k_f) \right]^{-1}, \quad (23)$$

$$\begin{aligned} \min_{F_c \in K_2} P(t, t_a) &= \exp \left[-t / \left(\min_{F_c \in K_2} T_0(t_a) \right) \right] \\ &= \exp \left\{ -n\lambda t \left[\frac{x\rho^m}{m!} \prod_{i=1}^m (n+i) + q^*(1+k_f) \right] \right\}, \end{aligned} \quad (24)$$

$$\begin{aligned} \max_{F_c \in K_2} P(t, t_a) &= \exp \left[-t / \left(\max_{F_c \in K_2} T_0(t_a) \right) \right] \\ &= \exp \left\{ -n\lambda t \left[\frac{x\rho^m}{m!} \prod_{i=1}^m (n+i) + q_*(1+k_f) \right] \right\}, \end{aligned} \quad (25)$$

where q_* and q^* – see Table 4; $\rho = \lambda \bar{I}_R$;

$$x = \begin{cases} \frac{\beta_m}{\beta_1^m}, & l = 1, \\ 1, & l \geq m. \end{cases} \quad (26)$$

Two-sided estimates of the average system recovery time are expressed (16) and (17), which should be substituted instead of Λ_1 the formula:

$$\Lambda_1 = n\lambda \left[\frac{x\rho^m}{m!} \prod_{i=1}^m (n+i) \right]. \quad (27)$$

To calculate the two-sided assessments of complex indicators of the reliability of the system $K_h(t_a)$ and $P_g(t, t_a)$ can be used (18)–(21), which should be substitutions, respectively (22)–(27).

In Table 6 the formulas for calculating two-sided assessments of system reliability indicators for an individual case $n = m = 1$ (duplicate system) are given.

Table 6. Duplicate system with time redundancy

Indicator	Limit values	Estimated formulas
$T_0(t_a)$	$\min_{F_c \in K_2} T_0(t_a)$	$(1/\lambda)[2\rho + q^*(1 + k_f)]^{-1}$
	$\max_{F_c \in K_2} T_0(t_a)$	$(1/\lambda)[2\rho + q_*(1 + k_f)]^{-1}$
$P(t, t_a)$	$\min_{F_c \in K_2} P(t, t_a)$	$\exp[-\lambda t(2\rho + q^*(1 + k_f))]$
	$\max_{F_c \in K_2} P(t, t_a)$	$\exp[-\lambda t(2\rho + q_*(1 + k_f))]$
$T_R(t_a)$	$\min_{F_c \in K_2} T_R(t_a)$	$(2\rho T_R^* + q^*(1 + k_f)\underline{T}_c^*(t_a))/(2\rho + q^*(1 + k_f))$
	$\max_{F_c \in K_2} T_R(t_a)$	$(2\rho T_R^* + q_*(1 + k_f)\bar{T}_c^*(t_a))/(2\rho + q_*(1 + k_f))$
$K_h(t_a)$	$\min_{F_c \in K_2} K_h(t_a)$	$[1 + \lambda(2\rho T_R^* + q_*(1 + k_f)\bar{T}_c^*(t_a))]^{-1}$
	$\max_{F_c \in K_2} K_h(t_a)$	$[1 + \lambda(2\rho T_R^* + q^*(1 + k_f)\underline{T}_c^*(t_a))]^{-1}$
$P_g(t, t_a)$	$\min_{F_c \in K_2} P_g(t, t_a)$	$\min_{F_c \in K_2} K_h(t_a) \exp[-\lambda t(2\rho + q^*(1 + k_f))]$
	$\max_{F_c \in K_2} P_g(t, t_a)$	$\max_{F_c \in K_2} K_h(t_a) \exp[-\lambda t(2\rho + q_*(1 + k_f))]$

Note: $T_R^* = \begin{cases} \frac{\beta_2}{2\beta_1}, & l = 1, \\ \frac{\beta_1}{2}, & l = 2; \end{cases}$ $\underline{T}_c^*(t_a)$ i $\bar{T}_c^*(t_a)$ - (14) and (15); q_* and q^* - Table 4; $\beta_1 = \bar{t}_R$; $\rho = \lambda \bar{t}_R$; $k_f = \lambda_f/\lambda$.

6 Two-Sided Estimates of the Reliability of TE Taking into Account Refusals and Failures in the Use of Time Redundancy

The purpose of this section is to construct two-sided assessments of reliability indicators, taking into account refusals and failures of the TE with time redundancy, which does not provide for the use of the structural reserve.

In the event of a failure, a number of operations related to the elimination of the consequences of the failure and aimed at restoring the efficiency of the TE. At the same time, the execution of these operations t_f is a random variable with a distribution function $F_f(t) = P\{t_f < t\}$ whose type is unknown and only two initial moments are defined

$$s_1 = \int_0^\infty x dF_f(x), \quad s_2 = \int_0^\infty x^2 dF_f(x), \quad 0 < s_1 < s_2 < \infty. \tag{28}$$

Refusal of the system (failure of functioning) arises at the moment of exceeding the time t_f of the permissible value t_a . Probability of this event

$$q = P\{t_f > t_a\} = 1 - F_f(t_a). \tag{29}$$

For the functional q the boundary values (lower q_* and upper q^* bounds) (Table 4) [23] are obtained, on the basis of which we obtain two-sided estimates of the system reliability:

$$\min_{F_f \in K_2} T_0(t_a) = (1/\lambda)(1 + q^*k_f)^{-1}, \quad (30)$$

$$\max_{F_f \in K_2} T_0(t_a) = (1/\lambda)(1 + q_*k_f)^{-1}, \quad (31)$$

$$\min_{F_f \in K_2} P(t, t_a) = \exp\left[-t / \min_{F_f \in K_2} T_0(t_a)\right] = \exp[-\lambda t(1 + q^*k_f)], \quad (32)$$

$$\max_{F_f \in K_2} P(t, t_a) = \exp\left[-t / \max_{F_f \in K_2} T_0(t_a)\right] = \exp[-\lambda t(1 + q_*k_f)]. \quad (33)$$

Integrated indexes of system reliability include a functional $I_3(F_c) = T_c^*(t_a)$ for which the lower $\underline{T}_c^*(t_a)$ and upper $\bar{T}_c^*(t_a)$ bounds (14) and (15) are obtained. On the basis of these boundary values, we obtain two-sided estimates of the mean time of system restoration, which are included in the formula of readiness:

$$\min_{F_f \in K_2} T_R(t_a) = (1 + q_*k_f)^{-1} [\bar{t}_R + q_*\underline{T}_c^*(t_a)], \quad (34)$$

$$\max_{F_f \in K_2} T_R(t_a) = (1 + q^*k_f)^{-1} [\bar{t}_R + q^*\bar{T}_c^*(t_a)]. \quad (35)$$

Formulas (18)–(21) can be used to calculate two-sided estimates, readiness coefficients $K_h(t_a)$ and operational readiness coefficients $P_g(t, t_a)$ in which the corresponding formulae (30)–(35) should be substituted.

7 Results of Complex Assessment of Reliability of TE with Limited Initial Information

The purpose of this section is to obtain two-sided assessments of the reliability of the TE of ICN with limited initial information. The given results allow quantifying the boundary values of the reliability values of the duplicate TE set with the following output data: we shall assume that the type of the distribution function of the connection time t_c of the backup device instead of the main rejected one is unknown, and only the first s_1 and second s_2 initial motives of this function are determined (2) which must satisfy the condition $0 < s_1^2 < s_2 < \infty$. The switching time is limited $t_c \leq Q$, that is, the switching time can not be longer than a certain fixed value Q . In this case, the inequality $0 < s_1^2 < s_2 < s_1 Q$ must be fulfilled. During making calculations we will take $s_1 = \bar{t}_c = 10 \text{ min.}$; $s_2 = 130 \text{ min}^2.$; $Q = 30 \text{ min.}$ and we use the boundary values of the functionals $I_1(F_c)$ and $I_3(F_c)$ (4) and (6), which are given in the Tables 1, 2, 3 and 4, as

well as formulas for determining the limit values of the reliability indicators of TE, which are obtained in Sects. 5, 6 [32, 34, 35].

In Tables 7 and 8 the results of the estimation of the marginal values of the reliability of the TE with structural, derating and time redundancy are shown (Table 7), as well as the time redundancy of TE (Table 8).

Table 7. Results of calculating the reliability of the TE with structural, load and time reservation

k_f	t_a , min.	Limit values of reliability indicators					
		$\min P(t, t_a)$	$\max P(t, t_a)$	$\min K_h(t_a)$	$\max K_h(t_a)$	$\min P_g(t, t_a)$	$\max P_g(t, t_a)$
0	4	0,9969	0,9986	0,999960	0,999988	0,99686	0,99858
	12	0,9976	~ 1	0,999637	0,999999	0,99724	0,999999
1	4	0,9938	0,9966	0,999937	0,999981	0,99374	0,99658
	12	0,9952	~ 1	0,999583	0,999999	0,9948	0,999999
5	4	0,9822	0,9902	0,999846	0,999954	0,9820	0,9901
	12	0,9856	~ 1	0,999442	0,999999	0,98510	0,999999

Table 8. Results of calculating the reliability of the TE with time redundancy

k_f	t_a , min.	Limit values of reliability indicators					
		$\min P(t, t_a)$	$\max P(t, t_a)$	$\min K_h(t_a)$	$\max K_h(t_a)$	$\min P_g(t, t_a)$	$\max P_g(t, t_a)$
1	4	0,994	0,9955	0,9961	0,9976	0,990	0,9931
	12	0,9946	0,997	0,9947	0,9983	0,9893	0,9953
5	4	0,9822	0,9891	0,9951	0,9981	0,9774	0,9872
	12	0,9851	0,9967	0,9863	0,9993	0,9716	0,9960

Analyzing the data of these tables and comparing them with the known data obtained in [15, 34] with full initial information, we see that the exact values of reliability indices are within the intervals limited by the obtained exact lower and upper estimates in all cases considered. This testifies to the adequacy of the proposed mathematical models of reliability of the TE, which allow obtaining two-sided assessments of the reliability of the individual TE of ICN under a priori uncertainty, in which different types of redundancy are provided.

8 Conclusions

In the article the calculated relations are obtained for construction of two-sided estimates of reliability of TE ICN, taking into account stable refusals and failures under a priori uncertainty. It has been shown that failures can significantly impair reliability even in the presence of structural and time redundancy. Therefore, when investigating the issues of ensuring the required level of reliability of TE, one of the important

directions should be the justification of effective ways to reduce the intensity of failures and their impact on the processes of ICN.

References

1. Netez, V.: Reliability of communication networks in IEC standards. *Commun. Bull.* (2), 13–15 (2014)
2. Ahmada, W., Pervez, U., Qadirb, J.: Reliability modeling and analysis of communication networks. *J. Netw. Comput. Appl.* **78**, 191–215 (2017). <https://doi.org/10.1016/j.jnca.2016.11.008>
3. Skulysh, M.A., Romanov, O.I., Globa, L.S., Husyeva, I.I.: Managing the process of servicing hybrid telecommunications services. *Quality Control and Interaction Procedure of Service Subsystems*. In: *Advances in Intelligent Systems and Computing*, vol. 889, pp. 244–256 (2019)
4. Semenov, A., Posyayeva, Z.: Increased reliability of optical communication systems. *Network Solutions Magazine/LAN* (2017). <https://www.osp.ru/telecom/2011/12/13012037>. Accessed 02 Jan 2017
5. Krendentser, B.: *Maintenance and reliability of systems with time redundancy: monograph*, p. 384. Phoenix, Kyiv (2016)
6. Netez, V.: Reliability of communication networks during transition to NGN. *Commun. Bull.* (9), 126–130 (2007)
7. Devis, V.: Dependability in the future battle network system - transport layer ability to maintain quality of service. Center for Applied Research, Estonian National Defense College, Tartu, Estonia, pp. 211–228 (2016)
8. Stepanova, I., Abdulvasay, A., Zhuhev, N.: Analysis of promising approaches to increase the reliability of converged corporate communications networks. *Commun. T-SOMM* **9**, 44–51 (2015)
9. Kononova, I.: Models of fail-safe telecommunication equipment of communication networks in different ways of structural redundancy. *Collection of scientific works*, no. 3, pp. 88–94. VITI, Kyiv (2017)
10. Meikshan, V.: Analysis of the influence of equipment failures on the operation of a multiservice network with adaptive routing. *Reports of the Academy of Sciences of the Higher School of the Russian Federation. Technical sciences*, vol. 2, no. 15, pp. 69–80 (2010)
11. Koval'kov, D.: Mathematical models for reliability evaluation of a multiservice access node. *Radio Eng. Telecommun. Syst.* **80**(2), 64–71 (2011)
12. Lemeshko, A., Kozlova, E., Drobot, O.: Mathematical model of fault-tolerant routing, representation by algebraic equations of co-state of MPLS-network. *Inf. Proces. Syst.* **2** (109), 217–220 (2013)
13. Popovsky, V., Volotka, V.: Methods of analysis of dynamic structures of infocommunication systems. *Telecommun. Inf. Technol.* **110**(3), 5–9 (2014)
14. Ignatov, A., Shuvalov, V.: The reliability of subscriber access networks LR-PON. *T-comm* **9** (5), 25–30 (2015)
15. Mogilevich, D., Krendentser, B., Minochkin, A.: *Estimation of operational and technical characteristics of telecommunication objects at a priori uncertainty: monograph*, p. 332. Phoenix, Kyiv (2012)
16. NSU 2860-94: Reliability of technology. Terms and definitions, Kiev (1995)

17. Mogilevich, D., Kredentser, B., Butochnov, O., Minochkin, A.: Reliability of systems with redundancy: methods, models, optimization: monograph, p. 342. Phoenix, Kyiv (2013)
18. Cisco IOS Release 12.2 (14) S. OSPF Update Packet-Pacing Configurable Timers. Cisco, (2010). http://www.cisco.com/c/en/us/td/docs/ios/12_2s/feature/guide/fsospfct.html. Accessed 01 May 2018
19. Li, T., Cole, B., Morton, P., Li, D.: RFC 2281. Cisco Hot Standby Router Protocol (HSRP). Network Worces Group (1998). <http://www.ietf.org/rfc/rfc2281.txt>. Accessed 01 May 2018
20. Tymoshenkov, S., Simonov, B., Goroshko, V.: Fundamentals of reliability theory: a textbook, p. 445. Yurait, Moscow (2015)
21. Viktorova, V., Stepanyans, A.: Models and methods of calculating the reliability of technical systems, p. 256. Lennand, Moscow (2016)
22. Kononova, I., Mogilevich, D., Kredentser, B.: Analytical model of complex assessment of reliability of the duplicate set of telecommunication equipment taking into account refusals and failures. Collection of scientific works, no. 4., pp. 48–56. VITI, Kyiv (2017)
23. Mogilevich, D., Kredentser, B., Vishnivsky, V.: Estimation of the reliability of reserved systems with limited initial information: monograph, p. 336. Phoenix, K. (2013)
24. Mogilevich, D., Kredentser, B., Minochkin, A.: Reliability of systems of occasional use in conditions of limited information about the initial data. Collection of scientific works of VITI NTUU “KPI”, no 2, pp. 37–41. VITI NTUU “KPI”, K. (2010)
25. Stoykova, L.: Necessary and sufficient condition for the extremum of the Lebesgue- Stilltiess integral on the distribution class. *Cybern. Syst. Anal.* (1), 72–75 (1990)
26. Ermoliev, Y., Kaniovsky, Y.: Asymptotic properties of some methods of stochastic programming with a constant step. *J. Comput. Math. Math. Phys.* **415**(2), 356–366 (1979)
27. Korolyuk, V.: Stochastic models of systems: a manual, p. 136. Lybid, Kyiv (1993)
28. Stoykova, L.: Some weak prerequisites for the extremum of a fractional-linear functional in the class of distribution functions. *Rep. Acad. Sci. Ukraine* **126**(12), 89–96 (1993)
29. Stoykova, L.: Investigation of generalized Chebyshev inequalities with application to mathematical reliability theory: author’s abstract. dis doctor of physical and mathematical sciences, 01 May, 01 INS-T cybernetics of NAS of Ukr. them VM Glushkov Kyiv, p. 31 (1995)
30. Stoykova, L.: Generalized Chebyshev inequalities and their application in the mathematical theory of reliability. *Cybern. Syst. Anal.* (3), 139–144 (2010)
31. Kovalenko, I., Stoykova, L.: Some extreme problems of reliability theory. In: Proceedings of the USSR Academy of Sciences. Technical Cybernetics, Issue 6, pp. 19–23 (1986)
32. Mogilevich, D., Kononova, I., Kredentser, B.: Analytical model of complex assessment of reliability of the duplicate set of telecommunication equipment taking into account failures and refusals. Collection of scientific works (WITI), no. 17, pp. 48–56. VITI, Kyiv (2017)
33. Mogilevich, D., Vishnivsky, V., Kredentser, B.: Methodology for determining the limit values of quality indicators of maintenance of systems of episodic use with a time redundancy with limited information on the distribution of work to refusal. Collection of scientific works of VITI NTUU “KPI”, no. 2, pp. 55–63. VITI NTUU “KPI”, Kyiv (2012)
34. Mogilevich, D., Kredentser, B., Vishnivsky, V.: Estimation of the reliability of reserved systems with limited initial information: monograph, p. 336. Phoenix, Kyiv (2013)
35. Kononova, I.: Analysis of the impact of failures on the reliability of telecommunication equipment of communication networks. Collection of scientific works, no. 3(34), pp. 56–67. NADPSU, Kh (2017)

Telecommunications



Combined Over-the-Horizon Communication Systems

M. Ilchenko^(✉), S. Kravchuk^(✉) , and M. Kaydenko^(✉) 

National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”, Peremoga Avenue 37,
Kyiv 03056, Ukraine
ilch@kpi.ua, {sakravchuk, kkk10}@ukr.net

Abstract. The chapter presents the scientific and technical principles of construction of new hybrid combined over-the-horizon communication systems with the use of reference small-sized stations of troposcatter communication, relaying intellectual aeroplatforms and artificial formations. These principles are based on the use of new technologies and software-defined and cognitive radio, cooperative relaying, machine-to-machine, effective interaction system, hardware and application protocol levels.

Keywords: Troposcatter communication · Over-the-horizon communication · Scintillation · Ray tracing · Relay aeroplatforms · Artificial formations

1 Introduction

Radio systems of the over-the-horizon (or the line of sight b-LoS - beyond-Line-of-Sight) communications (OTHC) are widely used, particularly in the military field (in strategic, operational and tactical levels of management), special services, rescue services and emergency situations. Particular attention is paid to the OTHC of the so-called troposcatter (short for tropospheric scatter) radio relay stations (TRRS), which are more difficult to implement than the microwave system line of sight. The principle of such tropospheric radio communication lines is based on two mechanisms for the propagation of radio waves: in the intervals of more than 65 km radio waves dissipate on inhomogeneities of the troposphere; at short intervals - diffraction distribution of radio waves. The main features of such tropospheric propagation of radio waves, related to the physical nature of this phenomenon, is a very large attenuation of the radio signal in the area of propagation and fading of the signal at the receiver input, which is characterized by a selective frequency.

Early tropospheric communication systems had the bandwidth of a radio channel sufficient to transmit voice and small data streams over long distances without intermediate repeaters. These systems were characterized by the use of more powerful power amplifiers, very large antenna mirrors, which were bulky for deployment and difficult to transport. The improvement of the efficiency of the modern new generation TRRSs is achieved by eliminating these deficiencies in them while maintaining the necessary noise immunity in conditions of multipath fading. However, the rapid

development of information networks requires support of the TRRS high bandwidth values within the range of 20...100 Mbps at a readiness rate of not less than 0.98, which is difficult to implement by traditional approaches.

The main obstacle to a significant improvement in the characteristics of the TRRS is a significant signal attenuation in the case of “distant tropospheric propagation,” about 80% of the transferred power is lost. The cardinal solution to this problem is proposed by using in the conditions of beyond-line-of-sight for the transmission of radio signals artificial movements in the troposphere or a plurality (constellation) aeroplatform drones that will perform the function of passive or active retransmission of the signal. Moreover, the use of relay nodes on the basis of aeroplatforms opens up new promising opportunities, in particular, the redistribution of traffic in various directions and the creation of local areas of coverage of broadband radio access systems (with access point on the airplane), the formation of a multi-level telecommunication network with self-organization, which to a large extent does not depend on ground infrastructure. These properties are very important for providing telecommunication services areas of natural disasters, hostilities, large water remote, mountainous or desert spaces.

At present, over-the-horizon communication systems have mainly been used to create point-to-point networks, in which TTRSs serve as final nodes or relay stations [1–3] as base nodes.

Heterogeneous combined over-the-horizon communications systems (COTHCS), proposed by the authors of this work, are based on small-scale tropospheric communication stations with the ability to work in indirect modes (tropospheric scattering or re-reflection from artificial entities in the atmosphere), and direct (using one or the constellation of relay aeroplatform-drones) visibility. The retransmitting airplanes form a self-organizing constellation (swarm) in a mixed group, which includes various unmanned aerial vehicles (UAV). Grouply coordinated application of UAV can significantly improve the performance of relay processes and the possibility of implementing multi-station broadband radio access. As artificial structures, various passive obstacles may be involved. The inclusion in the OTHC as a repeater aeroplatform, in contrast to satellites, can withstand a low delay in data transmission and have a small energy to generate radio paths significantly lower cost system.

For such combined systems there was a special direction of research and development associated with the creation of portable TPPs with low small-scale indicators and a sufficient level of efficiency of electrical characteristics, which is determined in providing a sufficiently high throughput while ensuring high immunity with limited energy and overall performance.

The purpose of this work is to develop the scientific and technical bases for the construction of new heterogeneous combined systems of over-the-horizon communication with the use of reference small-sized stations of tropospheric communication and relaying intellectual aeroplatforms and artificial formations that will be based on the application of new technologies of cognitive and programmable radio, co-operative relaying, M2M, digital signal processing, organization of efficient interaction of system, hardware and application protocol levels, and simplification of the percentage sous management and monitoring of structural elements, increasing their availability and reliability.

2 Related Works

Application of TTRS for more structured distributed network architecture practically unknown. There are developments that can be used to build hybrid distributed networks.

In [4], an approach was proposed to improve the quality and capacity of over-the-horizon communication radio channel by jointly operating on the principles of cognitive radio in one location of microwave ground and satellite radio relay communications and troposcatter stations. This approach is called hybrid cognitive troposcatter communication. However, the improvement in the efficiency of the radio link troposcatter in this scenario occurs due to the redistribution of information channels between different communication means, reducing the load on the radio channel troposcatter under adverse conditions of its operation.

In [5], the over-the-horizon communication system includes, in addition to the signal transmitter, an acoustic source located on the surface of the earth in the direction to be communicated. The acoustic source is able to provide periodic changes in atmospheric density that allows changing the field of tropospheric scatter parameters and thus modify the signal sent to the horizon intersection of a particular point of reception. However, the method described here for influencing the channel of a tropospheric transmission does not set itself the goal of increasing the efficiency of the channel and the possibility of creating a geographically distributed system on its basis.

Currently, communication lines limit the autonomy of unmanned aerial vehicles (UAVs). In [6], a multi-channel communication line for long distances is proposed for UAVs with high quality of service, moderate throughput and affordable cost. This line is deployed using a second unmanned aircraft as a communications repeater and provides bi-directional coverage for telemetry and telecontrol and a downlink with high bandwidth for video. Such an ultra-compact system can be installed on medium-sized unmanned aerial vehicles for air-to-air communications in order to ensure greater flight autonomy. This system is based on an active circular array of linearly polarized circular patch antennas that are selectively activated depending on the desired direction of the beam. Such a system can be effectively used as a separate domain in the proposed Combined Over-Horizon Communications Systems.

In [7] considered the possibility of building a wireless network self-organization based on UAVs, suggested core network architecture and main characteristics of the channel. In this case, the key design features of on-board equipment, as well as new possibilities of applications of such architecture are determined. Work [8] focuses on the main characteristics of UAV-based communication systems, describes alternative scenarios for network architectures to provide wireless access to broadband communications services. The scenario of the coexistence of such networks with terrestrial broadband radio access systems, including WiMAX, in one coverage area with high QoS values is considered. In [9] it is noted that the inclusion in the telecommunication network with the support of cognitive radio units on the basis of UAV allows the use of co-operative relaying for significant improvement of network characteristics: increasing bandwidth and reducing the power of radiation. In [10] presented the possibility of providing the services “Internet of things” IoT through UAV, proposed implementation

of telecommunications system architecture based aero platform based on the principles of IoT. [11] gives a detailed overview of M2M communications in the context of LTE-A (Long-Term Evolution-Advanced) mobile networks; presented end-to-end network architectures and reference models for M2M connections. In addition, the main ongoing efforts to standardize future M2M telecommunication networks and their services are considered. Attention is paid to multilayer and multidimensional distributed networks based on M2M, involvement in the construction of such a network of nodes based on mobile platforms and the need for long-distance communication lines.

In [12–14] for zone Antarctica deployment of a distributed wireless telecommunication system based on high-rise aero platform rapid deployment and basic ground TTRS is proposed. The system proposed for the creation of mobile broadband subscriber radio access local services.

Currently, IPv4 and IPv6 technology is used to build any packet network, which uses one endpoint interface in one communication session, even when the device has more than one interface. In [15], the MPT software library (multipath software tool) was proposed, which allows using several interfaces (and, accordingly, several paths) in one communication session. There is no need to modify the software of the application: it uses one logical interface for communication, and the MPT software dynamically maps (according to a specific map the set of available interfaces/paths) packets (sent by the application) to the available physical interfaces. The results of this work provide an important opportunity at the network level to form a distributed communication structure offered by Combined Over-Horizon Communications Systems.

At present, efforts have been intensified to improve the efficiency of data transmission through the tropospheric scattering radio channel, based on the achievements of modern theory of wireless systems. In [16], the hardware implementation of broadband communication in the troposcatter channel was presented using SCFDE (Single Carrier Frequency Domain Equalization). Here was simulated multipath and frequency-selective damped nature of the broadband channel of the troposcatter line. The simulation results showed that the proposed SCFDE-based system has a link margin of 3.2 dB with QPSK and 2.8 dB with 16 QAM compared to an OFDM-based system. It was found that the theoretically achievable maximum data transfer rate for both systems with a bandwidth of 20 MHz was 36.76 Mbps using 16 QAM modulation.

In article [17], the possibility of using the OWTTT (one-way troposcatter time transfer) during tropospheric scattering is investigated. The time delays of the propagation of tropospheric scattering are calculated on the basis of the Snell theorem and the refractive profile models. OWTTB (one-way troposcatter time broadcasting) method is proposed for synchronizing several nodes over a large area. It was found that OWTTT errors can reach about tens of nanoseconds for remote transmission of time and nanoseconds for a relatively short communication distance, for example, up to 120 km. In general, the OWTTT can satisfy the requirements of most remotely distributed radio systems over a large area, in particular, systems based on over-the-horizon communications. But how to build such a system here is considered.

Due to multipath propagation, the tropospheric scatter channel has a certain temporal dispersion, which significantly degrades the quality of communication due to

intersymbol interference. As a rule, the coherence band can be used to conveniently describe the characteristics of the multipath effect of the tropical scattering channel. In [18], the transfer function based on the scattering field was obtained, and the closed-form expression of the coherence bandwidth for tropospheric scatter links through the frequency correlation function was determined. These results provide important assessments of the system limits on coherence bandwidth for the Combined Over-Horizon Communications Systems being developed.

With an increase in the bit-rate and an increase in the multipath fading of the signal in the tropospheric scattering field, the use of OFDM technology was proposed in [19], which increases the spectrum utilization rate and resists fading in the choice of frequency. Models have been proposed that compare the bit error rate between the system with OFDM and the system in which OFDM is not used. The results of such simulations have shown that OFDM can effectively overcome multipath fading and improve communication quality with tropospheric dispersion.

For tropospheric communication systems it is important exerted influence of intersymbol interference for equalization in the frequency domain with a single carrier using space-time processing techniques MIMO (multiple input multiple output) MIMO-SC-FDE (single-carrier frequency-domain equalization with multiple input multiple output). Most of the known methods of equalization of communication channels cannot completely eliminate the harmful effects of inter-symbol interference. In [20], taking into account noise equalization disadvantages predictive NP (noise-predictive) based on MMSE (minimum mean square error) and RISIC (residual intersymbol interference cancellation) in the system with a single input single output SISO (single input single output), proposed a combination of both of these alignment schemes. After application of both schemes for MIMO systems, a new method has been introduced MMSE-NP-RISIC alignment for tropospheric communication system with MIMO-SC-FDE.

Currently, research on the creation of higher-efficiency equipment for over-the-horizon communications is relevant and this is confirmed by the existing publications and specific developments of companies [21–23]. In particular, TeleCommunication Systems Achieves Unprecedented Performance and Comtech Systems have developed TCS 3T, providing data rates of up to 50 Mbps at a distance of 160 km; ASC Signal Corporation introduced its MilSatCom TDRS development, and Raytheon developed the C-band for the BLOS-T (Beyond Line of Sight - Troposcatter Communications) [24, 25]. However, before the mentioned TRRS was not the task of creating on their basis portable small-sized systems.

The analysis of the results obtained by scientists and engineers in relation to the problem of establishing over-the-horizon systems allows us to expect that the research and development of systems with the involvement of the TRMS is new. Such studies will provide innovative technical solutions to improve existing communication technologies, new software, mathematical models, methods and algorithms.

3 Architecture of Heterogeneous Combined Over-the-Horizon Communication Systems with the Use of Reference Small-Sized Troposcatter Stations and Relaying Intellectual Aeroplatforms or Artificial Formations

The proposed over-the-horizon communication systems are based on a series of (from two or more) reference small troposcatter stations with the ability to work in tropospheric scattering regimes, re-reflection from artificial structures in the atmosphere, using one or the constellation of relaying airborne drones, and with support for multiple frequency bands and external network interfaces. The retransmitting airplanes form a self-organizing constellation (swarm) in mixed groups that include various purposes (communication, monitoring, exploration, search of objects, aerial photography, etc.) of the UAV, or as part of autonomously functioning but coordinated controlled UAVs.

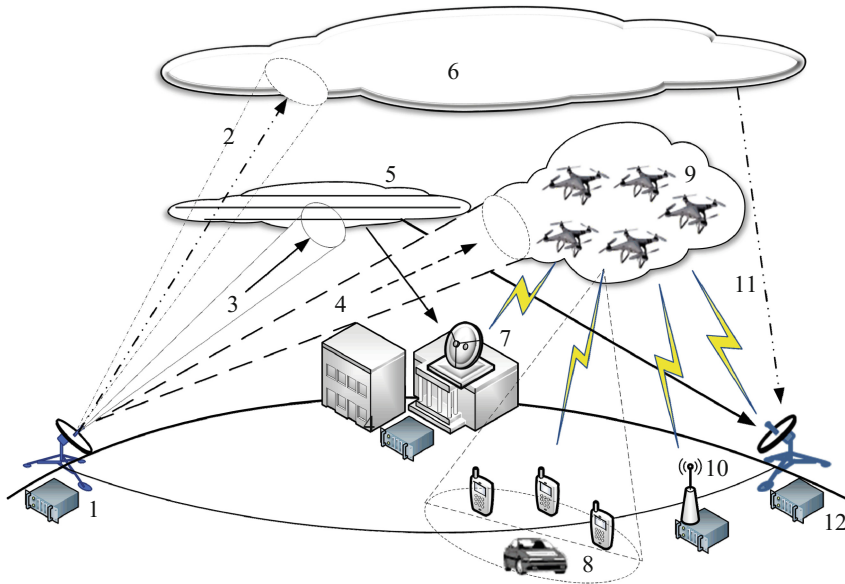


Fig. 1. Architecture of the heterogeneous over-the-horizon communications systems: 1, 12 - supporting modernized TRRS; 2, 3, 4 - radio beams in the directions of tropospheric scattering, artificial atmospheric formations and drone debris; 5 - layers of artificial formations in the atmosphere; 6 - layers of the atmosphere, where the formation of the volume of tropospheric scattering is possible; 7, 10 - separate radio equipment; 8 - radio terminals of mobile users; 9 - a layer of rows of drones; 11 - radiation created as a result of distant tropospheric propagation.

Grouply coordinated application of UAV in swarm can significantly improve the efficiency of relay processes and the possibility of implementing multi-station broadband radio access within the boundaries of such a constellation, while applying the technology of co-operative relay and inter-machine interaction M2M. Different passive

interferences can be used as artificial formations, which is the result of scattering of electromagnetic waves by foil, dipole, angular and lens reflectors, reflecting antenna arrays, ionized media and aerosol formations.

Architecture of heterogeneous over-the-horizon communication systems with the use of reference small-sized stations of troposcatter communication and relaying intellectual aeroplatforms and artificial formations is presented in Fig. 1. Such OTHC can operate in a point-to-point mode at distances from tens to several thousand km, or in combined modes “point-multipoint” and “multipoint-multipoint”. In the latter modes, with the help of relay nodes that are in the air, separate zones of the special user radio access with the provision of spatial, frequency and time separation can be formed. In this case, the OTHC can provide an extension, or the formation of a remote coverage area as part of a cellular network or broadband access of one operator. For example, in the case of the need to provide telecommunications services on a remote island or continent (research stations in Antarctica), or in the event of a natural disaster, anti-terrorist operations or combat clashes.

Next, consider the feasibility of implementing such an architecture and tools for its study.

4 Simulating the Operation of the Radio Link of a Over-the-Horizon Troposcatter Communication

Although the foundations of the theory of tropospheric scattering were developed in the last century, however, the development of over-the-horizon communications requires constant refinement of the known theoretical positions in accordance with new data on the nature of tropospheric scattering, atmospheric heterogeneities, the possibilities of new methods for calculating and estimating the radio waves propagation [26]. If earlier, to predict the propagation of radio waves during tropospheric scattering, only statistic methods of accumulation and analysis of experimental data and large-scale integrated models of tropospheric scattering volume were used, then the determined methods have become increasingly popular, especially on the basis of ray-tracing techniques [27–30]. When using this technique, the paths of rays can be traced from the transmitter to the receiver, taking into account the interaction with the surrounding structures (reflection, diffraction or transmission). Characteristics of the propagation of radio waves are predicted based on information about the rays, such as electric field strength, path length, angles of incidence and reflection. Here, there is a correlation between the accuracy of prediction and the required computing time. For high precision forecasting is necessary to trace the ray tracing of a higher order that undergo many interactions, while taking into account as many surrounding structures. This increases the number of intermediate processes and computing time. The main problem of this technique when tracking rays is acceleration of the calculation process.

So, in [31–34], using the technique of ray tracing, an analytical model of the ring scatter model (RSM) was developed that allows us to estimate the correlation of fading in tropospheric systems, depending on spatial, frequency and angular diversity. This model can use the results of real measurements of water and steam, taking into account air turbulence fluctuations. In addition, there was a comparative impact of the diversity

methods that are suitable for tropospheric communications, and various spatial-frequency techniques by obtaining their distribution of achievable data rates. However, the model uses the whole volume integral scattering cross-section of the differential scattering, determined using the Rayleigh scattering spectrum Kolmogorov (variation of refractive index) and defined for this volume of distribution of heterogeneities in the atmosphere. All this considerably limits the use of the model, especially for the study of the influence of electrophysical parameters and artificial formations on tropospheric scattering, and requires the development of a new, smaller-scale model of tropospheric scattering based on the technology of ray tracing, taking into account the scattering of radio waves on individual inhomogeneities and atmospheric formations.

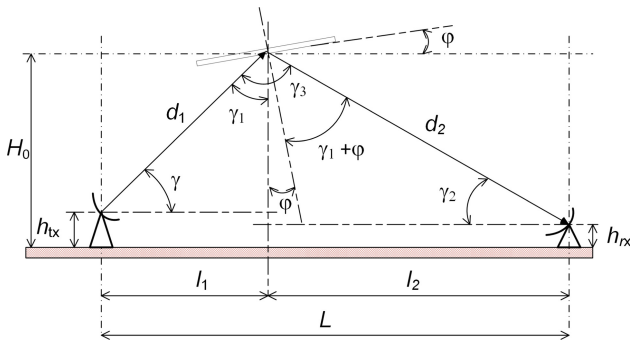


Fig. 2. The geometric model of the radio relay across a flat surface: γ - elevation angle of the axis of the main lobe of the antenna pattern of the transmitting station; γ_2 - elevation angle of the axis of the main lobe of the antenna pattern of the receiving station; γ_3 - the angle between the slope distances d_1 and d_2 ; φ - the angle of deflection between the perpendicular to the reflection area and the perpendicular of the projection of the high point on the floor of the h_{tx} and h_{rx} - the position above the surface of the antenna of the transmitting and receiving station, respectively; l_1 and l_2 are the distances between the projection on the surface of the earth of the reflection height point and the transmitting and receiving stations respectively; d_1 and d_2 are the sloping distances between the antenna of the transmitting station and the refraction point and between the antenna of the receiving station and the refraction point, respectively; L - total distance between ground stations; H_0 - height of the position of the height point of reflection relative to the earth's surface.

Therefore, the author developed a model of the radio link of the over-the-horizon communication to study the possibility of regulating (improving) attenuation along such a path by changing the electrophysical characteristics of the environment, in particular atmospheric turbulences, or the use of artificial formations in the atmosphere. On the basis of the developed model, investigate the specific laws of tropospheric scattering on inhomogeneities and artificial atmospheric formations in which the electrical parameters can be regulated and to determine the limits of the possible application of artificial passive noise in the troposphere for the directed reflection of the radio signal of the ground station with the possibility of creating an effective relay radio line of great length.

We will first introduce the simplest geometric model of the radio relay with a relay through a flat surface, which is used for simplified refinement calculations and presented in Fig. 2.

The losses of radio waves propagation with relaying through a passive surface area Area are: $path_loss = pl_{fs} - G_{passiv}$, where $pl_{fs} = 20 \lg (4\pi \times (d_1 + d_2)/\lambda_0)$ - loss in free space; λ_0 - wavelength in free space; $G_{passiv} = 20 \lg (4\pi \times Area \times \cos(\alpha) \times KPD/\lambda_0/\lambda_0)$ - passive relay transmitter ratio with area Area and coefficient of effectiveness of the surface repeater use KPD; $\alpha = \gamma_3/2$ - half the angle between the directions to the repeater and the terminal station. Geometric relations between the parameters presented in Fig. 2 the following:

$$\begin{array}{l}
 d_1 = (H_0 - h_{rx})/\cos(\gamma); \\
 d_2 = (l_2 \times l_2 + (H_0 - h_{rx}) \times (H_0 - h_{rx}))^{(1/2)}; \\
 l_1 = (H_0 - h_{rx})/\tan(\gamma); \\
 l_2 = L - l_1;
 \end{array}
 \left| \begin{array}{l}
 \gamma_1 = \pi/2 - \gamma; \\
 \gamma_2 = \text{atan}((H_0 - h_{rx})/l_2); \\
 \gamma_3 = \gamma_1 + \pi/2 - \gamma_2; \\
 \varphi = \gamma_3/2 - \gamma_1.
 \end{array} \right.$$

In the proposed model, the volume of tropospheric scattering is presented in the form of a multilayer structure, limited by the size of atmospheric turbulence or heterogeneity. According to the modeling of the layers, there are two variants of the representation of the propagation of radio waves through the trace of tropospheric scattering: (1) simplified - each individual layer is homogeneous with single electrophysical parameters; (2) complete - each layer is a complex structure of blocks of spherical diffusers with different electrophysical parameters.

In a simplified version of the Wolf-Bragg model, which is observed in X-ray diffraction in crystalline bodies, at the given angle of scattering of the tropospheric path θ_{rez} (that is, the traces of a certain length) and the wavelength λ effective scattering in the direction towards the point of admission is created only by those diffusers, which are separated from each other vertically by a distance $L = \lambda/2\sin(\theta_{rez}/2)$. The distance L between the parallel idealized layers of the atmosphere, where the sources of dispersion are located, equal to the size of atmospheric turbulence or heterogeneity. It is a spatial period of repetition of inhomogeneities in the troposphere. In a turbulent atmosphere, swirling is of a different size. The additional phase shift between such layers will be $\Delta\varphi = 4\pi L \sin(\theta_{rez}/2)/\lambda$.

Version of the full model provides full tracing all rays that can get to the plane 5 (Fig. 3).

Each layer of the atmosphere irregularities represented as independent block structures (*A* and *B*) with spherical diffusers [35] whose electrophysical parameters are described by the complex dielectric constant ϵ and the conductivity σ (Fig. 4). Block structures can be shifted to each other by δ_2 , and spherical diffusers form a quadrature grid with a displacement δ_1 . The formation of the tree of the trace occurs at several sublevels according to the structures *A* and *B* itself - large-scale and small-scale. The technique of ray tracing used is similar to [36–39].

The limiting conditions for further tracing of the ray in the volume of scattering (that is, its exclusion from the calculations) are:

- the instantaneous value of the power of the electromagnetic wave at the point of arrival of the ray P_M , taking into account the value of losses in the free space g_M , calculated from the point of arrival of the beam to the position of the antenna of the receiver with the gain coefficient G_{HPM} in the direction to the point of arrival of the beam, less sensitivity of the receiver of the ground station P_q , that is, $P_M G_{\text{HPM}}/g_M < P_q$;
- the additional phase shift of the beam exceeds the limiting phase shift, which corresponds to the coherent bandwidth of the receiving channel;
- the direction (vector) of the beam on the boundary loops of the volume has a direction other than the direction to the antenna direction diagram of the receiver of the ground station (this condition can be disconnected).

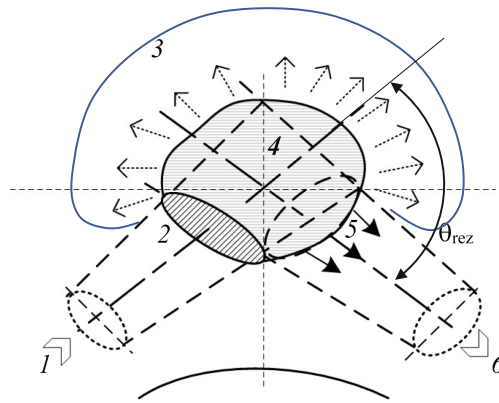


Fig. 3. Presentation tropospheric scattering volume 1 - radiation pattern of the transmitter antenna; 2 - the cross-sectional area of the pattern of the transmitter antenna with a dispersion volume; 3 - rays scattered in directions that do not coincide with the radiation pattern of the receiver; 4 - volume scattering; 5 - the cross-sectional area of pattern of the receiver antenna with the dispersion volume; 6 - pattern of the receiver antenna.

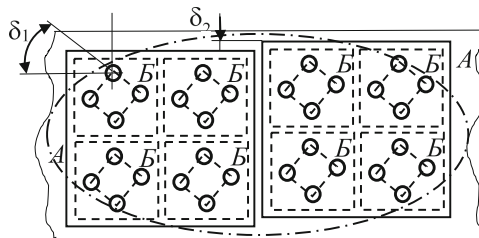


Fig. 4. Representation of a single layer of heterogeneities of the atmosphere in the form of independent block structures with spherical diffusers.

The following should be noted.

Any model describing the propagation through the Earth's troposphere will require a certain description of the state of the propagation medium. The simplest model of propagation can simply take on the global average atmosphere (for example, the coefficient of the Earth's globe $4/3$ for tropospheric propagation). But global and time variations may be large, and any model that aims to take into account this variability will be needed to represent the variability of meteorological data.

The ITU-R organization has developed a number of methods for this. Global variability is usually described by maps, and for the troposphere, the main references of ITU-R are Recommendation P.453 for the effects of the refractive index and Recommendation P.837 for rain effects. Digital versions of these maps are available on the website of the 3 ITU-R Research Group. Recommendation P.453 provides maps of different gradient and current statistics that were derived from measurements of radiosonde and models of multiple weather conditions. Some data on seasonal differences and time variability of the surface layer gradients are given.

In some cases, the method uses a combination of maps and simulations to provide the necessary data. For example, the calculation of rain deterioration in ITU-R P.837 takes three parameters that are geographically determined using maps, and uses them in formulas that determine the temporal dependence of rain relief. In other cases, a simple formula was found, sufficient to express the geographic variability of the meteorological parameter.

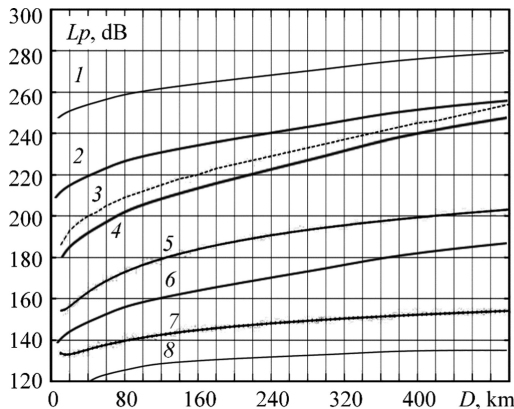


Fig. 5. Dependence of median propagation losses of radio waves L_p by the distance between ground stations D : 1 - model of the complete version with the introduction of ϵ spheres of the imaginary component $\epsilon'' = 0.035$; 2 - model of the simplified version; 3 - tropospheric scattering support; 4 - model of the full version; 5 - model of the complete version with the dilution of the dispersion volume in the first three of its layers (in block A, only one block B with a proportional scaling of the spheres); 6 - model of the full version with an increase of ϵ spheres; 7 - loss of distribution in a free space of length $2d$; 8 - model of the full version with the introduction into the conduction fields $\sigma = 10^{-3}$ S/m (water vapor with smog) of area 10 m^2 .

The calculations were made for a radio path with the following characteristics: the height above the Earth's surface of the lower point of the dispersion volume (the appearance of turbulence and heterogeneity) is 8 km; the distance from the ground stations to the lower point of the scattering volume is the same as their elevation angles (the distance diagonally to the point is $d = d_1 = d_2$), the operating frequency of the radio waves is 4.7 GHz, the height of the location of the antenna stations above the Earth's surface is 5 m, the diameter of the mirrors of the parabolic antennas 1.8 m. The accepted model of the planar surface of the Earth, the conditions of the standard atmosphere, atmospheric turbulence of the Middle European latitudes of the summer period (turbulence sizes up to 500 m), angles of the spatial structure of the scattering volume with respect to the plane of the horizon varied to 5° . For spheres of scattering as a material, water was received, and the distribution of the number of drops per cubic meter of air was determined by the gamma-distribution of their sizes.

Calculations of supporting median losses on the tropospheric scattering line were carried out in accordance with [26]. Support losses were used to conduct a comparative analysis of the correspondence of the results obtained with the proposed model with the reference loss data. Also, according to the reference losses, the calibration of the model was performed because it has certain components of simulation (selection δ_1 , δ_2 , ε , spheres size, followings of blocks *A* and *B*, height and shape of turbulence, required number of branches of tree of ray tracing, restriction of iterations in the layer of heterogeneities of the atmosphere and the angles of inclination of such a layer to the horizontal plane).

The simulation results using the proposed loss model on the radio path are shown in Fig. 5. The obtained results testify to the adequacy of the proposed model with the actual experimentally obtained losses during tropospheric scattering (curves 2, 3, 4). The introduction of additional attenuation ε'' to scattering fields causes almost linear increase in the loss of tropospheric scattering without changing the slope of the loss curve (curve 1). The growth of the real part ε leads to a decrease in losses (curve 6), but to a certain value ε_{lim} , after which the losses begin to grow quite fast. The value of ε_{lim} depends on many factors, for example, all the angles of inclination of the structures of spheres and layers, the angles of the location of antennas of ground stations, the length and height of turbulence. The great influence on the loss of tropospheric scattering results in the combination of blocks *A* and *B* with a proportional scaling of the position of the spheres (especially their inclination) (curve 5). Such a combination may simulate virtually any atmospheric turbulence. Introduction to the conduction fields dramatically reduces the loss on the radio path (curve 8), with the greater the conductivity, the smaller the layers of the dispersion volume is involved in the formation of radio paths. This is due to the fact that the lower layers are practically shielding the upper, reaching the effect of passive relay from a plane that has conductivity. Practically here the process of reflection from artificial reflectors is simulated: dipoles made of aluminum foil; angle reflectors, consisting of mutually perpendicular faces, whose dimensions considerably exceed the length of the incident radio wave; Lüneberg lenses, representing a dielectric layer with a variable refractive index; Van Atta array.

5 Supported Small-Dimensional Troposcatter Station

For heterogeneous combined over-the-horizon communication systems, the small TRRS 464465.001 has been developed as the reference TRRS. This station is created on modular principle based on a single uniform designs developed using modern microelectronic components and software-definable systems.

The main components are the following stations (Fig. 6): antenna-rotating device (ARD); transceiver unit (TRU); modem device (MD); power supply; cables (high-frequency, power supply and signaling); remote control and control device (mobile gadget).



Fig. 6. Compact TRRS: 1 - modem device; 2 - antenna-rotating device; 3 - transceiver unit; 4 - cables; 5 - power supply.

The compact TRRS 464465.001 has the following technical characteristics:

- operating frequency range 4.4–5.0 GHz;
- maximum speed of receiving/transmitting digital information stream up to 10 Mbps;
- Information interface 10/100/1000 Base-T, available ports with possibility of encapsulation of the external flow of converter E1 (G.703) in Ethernet;
- protocol and interface for data transmission IP (TCP/IP), Ethernet;
- range radio link with the availability of no worse than 0.98, and bit error rate no worse than 10^{-6} in digital flow rate: 512 kbps up to 130 km and 2,048 Mbps–90 km;
- the time of deployment and entry into the connection of the TRRS is no more than 15 min;
- automatic discrete (in 1 dB increments) output power regulation - not less than 30 dB;
- Integrated system of functional control without radiation in ether;
- the management of TPNS is carried out remotely at a distance of not more than 100 m.

6 Multilevel Adaptation to Change the Conditions of the Station

Terms of TRRS differ significantly from the conditions of microwave station line of sight. In the tropospheric channel there are complex fading (frequency-selective, fast and slow fading), which are daily and seasonal. To compensate for disturbances that occur in addition to the traditional methods of dealing with them (high power transmitter and channel equalizers use) in this TRRS applied multilevel adaptation to changing conditions of the station.

The station uses four levels of adaptation: frequency, modulation and encoding, spatial (in two-channel spatial diversity mode) and MAC level.

The task of frequency adaptation is the choice of the optimal bandwidth in the frequency range of the station in the presence of prolonged frequency-selective fade (idle duration exceeds the duration of several frames) and noise (structured and unstructured). Two stages of adaptation are used: initial frequency selection and adaptation in the working mode of the station.

The estimation of the communication channel state is carried out by analyzing the amplitude-frequency characteristic (AFC) through the transmission path (rough estimate) in the selected frequency band of maximum width. At least a rough assessment of selected subject to continuous frequency-selective fading bandwidth, which can be less than the maximum. A rough estimation is performed by scanning during initialization and optionally select a new frequency band in operation. For accurate estimation of response in addition to analyzing the analysis of energy performance and noise immunity basic channel profile transfer. Selects the band with maximum power reserve and maximum throughput. Frequency adaptation in the operating mode is carried out in the presence of more than one channel, which can work as an independent one.

The initial initialization algorithm involves scanning a range of frequencies with frames of minimum length consisting of three frames of the preamble, one of which is used for synchronization, and two others estimating the frequency response of the path, after transferring the scan frame, pauses the length of one frame to switch to another carrier frequency. Frequency band used at initial scan of 7 MHz. The scan results determine the frequency bands in which the fade is minimal. In defined frequency bands, the maximum possible width is repeatedly scanned by frames of minimum length (1 ms), which include a preamble and pseudorandom data packets to assess the energy characteristics (SNR signal/noise ratio) and noise immunity. To scan using a profile with BPSK (Binary Phase-Shift Keying) modulation and maximum noise immunity correcting ability of the code. The noise immunity is determined by the FEC (Forward Error Correction) of the decoder. The band with the maximum SNR and minimum BER is considered optimal. At initial initialization, simultaneously with the choice of optimal band, frequency bands are excluded, in which there are signals from other radio equipment, including other troposcatter communication stations.

The algorithm of frequency adaptation in the main mode differs from the initial initialization mode in that the evaluation of the frequency response AFC of the entire frequency range is not carried out, SNR and BER are evaluated using base duration

frames (5 ms). The choice of the required band of working frequencies is carried out according to the results of the algorithm of initial initialization. Since the fading in the tropospheric channel are including the daily nature of the information about the frequency response in the range of operating frequencies updated with a periodicity of not less than 6 h.

The advantage of using the frequency adaptation mode in the station and the ability to set the grid frequency step with an accuracy of less than 1 kHz allows you to operate without prior frequency planning and “hard” fastening of the working frequency grid.

Adaptive modulation and coding (AMC) is performed only after frequency adaptation; assessment of the channel obtained by AMC used as required for accurate frequency adjustment. Adaptive modulation and encoding relate to methods in which several channel parameters can vary at the same time, while both modulation parameters and interference-encoding parameters are changed simultaneously, as well as power adaptation.

As the channel status indicator for adaptive modulation and encoding, the average signal-to-noise ratio per interval of one data frame and its statistical characteristics measured by the receiver is used. In this case, the physical layer parameters (modulation and noise immunity encoding) are grouped in the transmission profile, each of which determines the capacity and security. Five transmission profiles are used with BPSK modulation, coding speed 1/2 and Barker code extension, and four others use QPSK modulation at approximately coding speeds 1/2, 2/3, 3/4, 7/8.

Procedure using SNR as follows:

- (1) receiver evaluates the channel state information by measuring SNR in estimation window at the interval, which is the length of the frame;
- (2) information about the SNR, which takes into account all kinds of distortions, is converted into BER information for each potential mode assuming that the channel was AWGN (Additive white Gaussian noise);
- (3) for a certain BER elected regime that is optimal according to selected criteria (maximum capacity, the maximum transmission reliability, maximum energy efficiency);
- (4) information on the selected mode is transmitted to the transmitter;
- (5) the transmitter changes the profile to the optimal.

As an algorithm for adaptive modulation and coding using the original linear modified algorithm based on fuzzy inference. Using this algorithm has allowed to increase the bandwidth of the station by an average of 15 ... 20% in comparison with using the linear algorithm of adaptive modulation and coding.

Spatial adaptation is choosing the optimal configuration settings antenna system (azimuth and elevation), it is possible only if two-antenna system antennas with independent and automatic adjustable angles, otherwise it is only in the initial choice of the angles. Spatial adaptation is carried out independently of other types of adaptation by using each spatial channel as an independent, with the total bandwidth of the station increasing by a factor of two.

Adaptation to the MAC level is carried out regardless of other types of adaptation and is to retransmit the incorrectly accepted data packet. The maximum number of

retransmissions of a packet is limited to the maximum allowable latency of the packet, which is a regulated parameter and does not exceed 500 ms.

6.1 Transceiver Unit

The need to create a compact, small-scale TRU, which at the same time should be quite powerful (having high levels of power for radiation), has led to the maximum integration of all the functional units of the TRU into a single technological design using the most advanced elemental base and the latest materials. Such a design of the TRU consists of two functionally interconnected modules: the base active duplex module (BAD) and the modulus of the up and down converter (UDC) with the control system of the control panel.

The main characteristics of the TRU are as follows:

- range of operating frequencies, MHz, in subbands: lower - 4435 ... 4555, the upper - 4705 ... 4825;
- maximum output power of the transmitter on the antenna flange, W - 140;
- the level of the output signal of the IF L-band at receiver, mW - 0,2 ... 0,5;
- depth of automatic control of gain of the receiver, dB - 35;
- relative level of the side components of the unmodulated spectra of the input signal to the transmission and output signal at the receiver, dBc - minus 67;
- receiving channel noise ratio relative to the antenna flange, dB - 3.2.



Fig. 7. Appearance of the TRU.

Appearance of the TRU is shown in Fig. 7.

6.2 Modem Device

Modem device provides modulation, demodulation, formation of information, voice and service data streams on radio and wire interfaces, multilevel adaptation of the station to constantly changing working conditions, synchronization, choice of operating modes of the station, control of the station and its individual units, monitoring station status, etc.

The modem device is a full-featured dual-channel modem module built on the basis of advanced technology of SDR (Software-Defined Radio) in conjunction with SoC chip (System-on-a-Chip). This variant of the modem device is oriented to use in stations, not only with one-channel mode of operation, but also with a dual-channel mode with polarization or spatial diversity. The appearance of the modem device is shown in Fig. 8.

The sharing of SDR and SoC technologies in the creation of the modem equipment of the tropospheric station provided the following benefits:

- the possibility of further upgrading without changing the hardware platform by using multiple program profiles to download the configuration of the station depending on the selected mode of operation;
- the use of high intermediate frequency (IF) in comparison with traditional solutions, the formation of the working frequency grid and the choice of operating frequency directly in the modem, rather than in the transmitter unit, which allows to increase the speed of frequency-frequency switching in the application of the frequency adaptation mode;



Fig. 8. Appearance of the modem device (right to the left: WTY0122 type connector with the button “Power 220 V”, DN 3 - control, N type - IF signal RX1, type N - IF signal TX1, N type - IF signal RX2, Type N - TX signal TX).

- filtration using a combination of programmed reconfigurable digital and analogue filtering of the main selection, which allowed the creation of a modem device that can be used over a wide frequency range limited by frequency dependent control elements (maximum range from 70 MHz to 6 GHz);
- a compromise distribution of resources for implementation of high-speed and low-speed processes of digital signal processing, decision-making algorithms and control algorithms between processor cores and FPGA (field-programmable gate array);
- implementation of the frequency adaptation mode (choice of frequency carrier and bandwidth) with rapid change of frequency carrier and bandwidth due to the use of digital frequency synthesizers and a set of programmed controlled analog and digital filters of the basic selection with specified profiles of adjustment of coefficients;

- implementation of the adaptive modulation and encoding mode with a quick change of profile implemented on the FPGA, and the algorithm for making a solution implemented on an ARM (Advanced RISC Machine) -type processor;
- implementation of MAC-levels of wired and wireless interfaces was implemented on one high-performance crystal, which allowed providing the necessary performance;
- protocols of the work of the modem equipment and the station were programmed in high-level languages, which allowed shortening the development of equipment;
- operation of the modem equipment and the station is carried out using a single operating system based on Linux.

6.3 Configuring the TRRS System

TRRS system can be configured for diversity reception (transmission) or space-time multiplexing MIMO, as shown in Fig. 9. The readiness coefficient of communication in the transmission of data at a speed of 512 kbps, depending on the distance between the TRRS is presented in Fig. 10. Variants of configuration of the equipment of the TRRS (various variants of the station’s configuration) are presented in Fig. 11. The scheme of implementation of increasing the speed of the system through the independent multiplexing of streams of digital data is shown in Fig. 12.

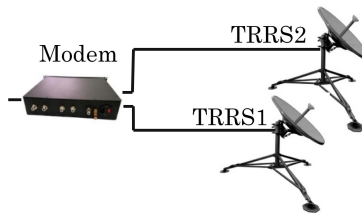


Fig. 9. TRRS configuration for diversity reception/transmission.

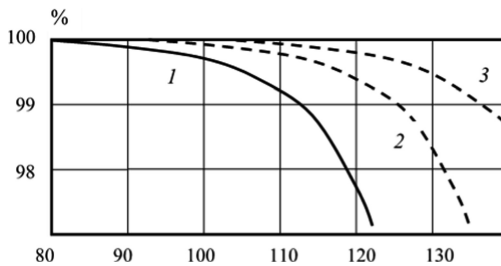


Fig. 10. A figure caption is always placed below the illustration. Short captions are centered, while long ones are justified. The macro button chooses the correct format automatically.

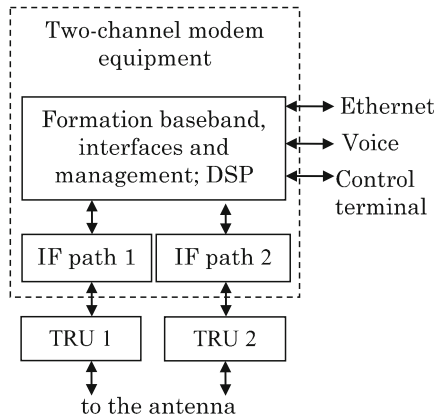


Fig. 11. Formation of the system with diversity.

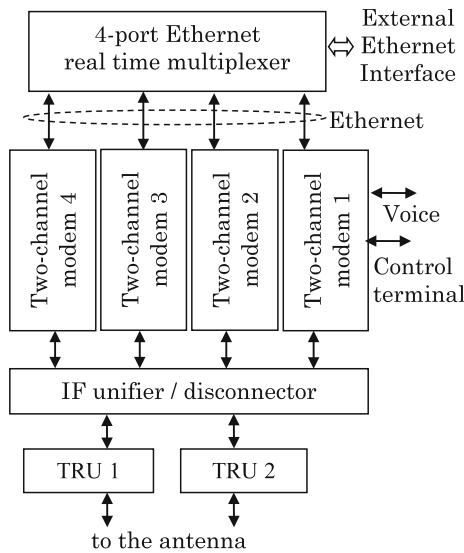


Fig. 12. The scheme of realization increasing rate system.

7 The Level of the Use of Airplane Platforms

As air platforms for COHCS, various flying unmanned vehicles of two categories can be used: individual heavy drones of long flight, small (even disposable) gliders and drones. On the basis of the second category UAVs, a swarm can be formed, having functions of both external control and self-organization. At the same time, they can form special antenna arrays or any reflection surfaces of radio waves with the help of

their individual antenna elements (Fig. 13). Depending on the angle of inclination of the antenna elements on the drone (or the inclination of the drone itself), one or another service area will be formed on the surface of the earth. At the same time, these drones are used for the implementation of both relay distribution and screening applications of the combined COTHCS wireless system.

The first category of power UAV have a powerful transceiver and antenna system, which supports radio links with supporting TRRSs and distributed over the unmanned aerial platform with terrestrial or aerial wireless terminals. They can also serve as coordinating hubs for drones and the COTHCS system itself, or a 5th generation mobile communication relay station with an Internet-based support network and M2M support function.

Structurally functional principles of the construction of the onboard control and communication system take into account that the system should ensure continuous data exchange on the control channel and telemetry with high reliability, as well as continuous transmission of data through the data link towards a ground-based system with guaranteed quality. The general structure of the onboard system of control and communication of the UAV of the first category is shown in Fig. 14.

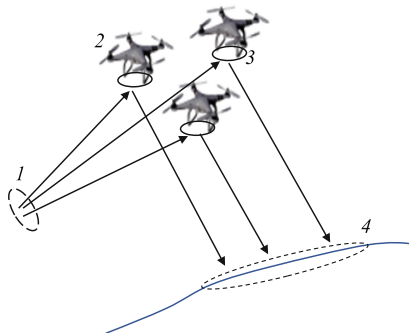


Fig. 13. Example of service swarm formation zone drones: 1 - supporting beam radiation pattern TRRS; 2 - quadcopter (drone); 3 - an antenna element in the form of a special reflective surface; 4 - area of service on the Earth's surface.

The control system consists of: control channel and telemetry with an antenna of the corresponding type; data transmission channels; antenna system with automatic maintenance functions; CPU board control and data processing; switching and routing unit; a set of cameras for video surveillance and flight control; flight controls and determine the orientation of the UAV space.

The control channel and telemetry are used to transfer UAV control data from the ground station and telemetry data from the aircraft to the ground station. To ensure continuity of management of the aircraft at all stages of flight using an antenna with circular radiation pattern. For the control channel, a frequency range of 30...470 (510) MHz can be used. The data rate in the channel is 19.2 ... 38.4 kbps and can be increased up to 100 kbps if necessary.

For data transmission channels are used in higher frequency bands of 700, 800, 900 MHz, 2.4, 5.8, 27, 42 GHz. At the same time two channels in two different ranges can be used at the same time to ensure the guaranteed quality of data transmission in conditions of deliberate interference. Data channels are asymmetric with time duplex, the channel towards the ground station is high-speed, the line down/line up can be 8/1, 16/1. Modulation in transmission channels - multi-position BPSK, QPSK, KAM 16, KAM 64, KAM 256. Interference-encoding - cascading code, block turbo code or LDPC code with variable coding speed. To ensure the fight against unintentional interferences uses adaptive modulation and coding. Due to the fact that data channels are high-speed and use a deliberate control mechanism, these channels can be used additionally as duplicates for the control channel and telemetry, or for the transmission of control and telemetry data simultaneously on three different routes.

In order to determine the orientation of the UAV in space, the satellite navigation systems (GPS, Galileo, Glonass) is used to determine the coordinates of the UAV and, as a duplicate, the WPINS (without platform inertial navigation system). Despite the fact that the accuracy of the WPINS on the basis of MEMS sensors is relatively low, it can be used to provide an orientation of the UAV in the absence (deliberate suppression) of GPS signals.

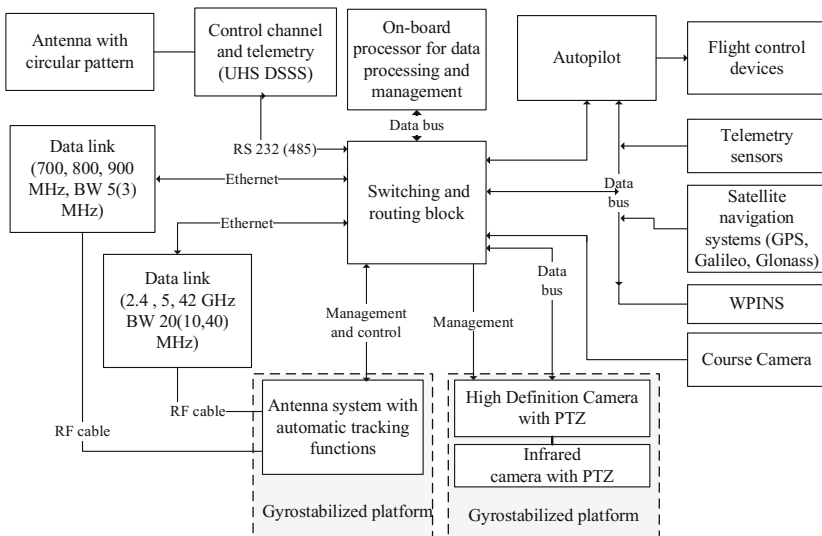


Fig. 14. General structure of the UAV control and communication system.

The on-board data processing and control unit, and the switching and routing unit, are organized on the basis of System-on-a-Chip system-to-crystal system technology, which combines the Hard Processor System (HPS) and FPGA (Field- Programmable Gate Array).

The switching and routing block are implemented on the FPGA SoC and operates under the control of the onboard processor for data management and data processing.

FPGA features functions such as packet generation, high-speed routing and switching, support for hardware interfaces and data bus with different speeds and different bit rates. In addition, some functions that require high-speed computing, as well as functions for forming control signals for gyro-stabilized platforms can be carried over.

8 Conclusions

The scientific and technical principles of construction of new heterogeneous combined over-the-horizon communications systems with the use of reference small-sized stations of tropospheric communication and relaying intellectual aeroplatforms and artificial formations, which will be based on the application of new technologies of cognitive and program-defined radio, co-operative relay, M2M, digital processing signals, organization of effective interaction of system, hardware and application protocol levels, and simplification of the control and monitoring process anchor elements, increase their availability and reliability.

In particular, the results of the development of a modern portable tropospheric radio relay station in the range 4.4 ... 5.0 GHz as general and special purpose (dual purpose) are presented. The station is characterized by low power consumption, compact dimensions and low mass due to the use of the latest elemental base of microelectronics in an extremely high frequency range, modern radio technologies for the formation and transmission of the signal through the tropospheric channel.

To compensate for the obstacles that arise in this TRRS, multilevel adaptation has been applied to change the external conditions of the station. It uses four levels of adaptation: frequency, modulation and encoding, spatial and MAC level. This allows the radio channel to reach the maximum speed of receiving/transmitting digital information stream up to 10 Mbps (protocol and data interface - IP (TCP/IP) and Ethernet)

The sharing of SDR and SoC technologies in the creation of modular equipment at the tropospheric station provides the opportunity to further upgrade the station without changing the hardware platform by using multiple program profiles to download the station configuration depending on the selected mode of operation.

The model of losses on the radio link of the horizon tropospheric connection (distant tropospheric propagation or tropospheric scattering), based on the technique of ray tracing in two variants: simplified with homogeneous layers and complete with combination of blocks of scattering sphere structures in each of the layers, was developed.

The study of the possibility of improving the attenuation along the trajectory of tropospheric scattering by controlling the change in the electrophysical parameters of inhomogeneities and artificial formations in the atmosphere was carried out.

The possibility of using artificial passive formations in the troposphere for the directed reflection of the radio signal of the ground station with the possibility of creating an effective relay radio line of large length is presented.

References

1. Freeman, R.L.: Radio System Design for Telecommunications, 3rd edn, p. 880. Wiley, Hoboken (2007)
2. Roda, G.: Troposcatter Radio Links. Artech House Publishers, Boston (1988)
3. Ilchenko, M.Y., Kravchuk, S.O.: Telecommunication Systems, p. 730. Naukova dumka, Kiev (2017)
4. Li, C., Chen, X., Liu, X.: Cognitive tropospheric scatter communication. *IEEE Trans. Veh. Technol.* **67**(2), 1482–1491 (2018)
5. Nemit, J.T.: Over the horizon communication system. US5017923A Patent Number: 5,017,923, 21 May 1991
6. Mu, Y., Blázquez García, R., Aguares Palomar, J., Sun, X., García Tejero, A., Fernández González, J.M., Burgos García, M., Sierra Castañer, M.: Over-the-horizon communications system for UAVs based on intelligent antennas. In: XXXI Simposium Nacional de la Unión Científica Internacional de Radio (URSI 2016) Madrid, Spain, pp. 1–4 (2016). <http://oa.upm.es/46490/>
7. Zeng, Y., Zhang, R., Lim, T.J.: Wireless communications with unmanned aerial vehicles: opportunities and challenges. *IEEE Commun. Mag.* **54**(5), 36–42 (2016)
8. Alsamhi, S.H., Rajput, N.S.: An intelligent HAP for broadband wireless communications: developments, QoS and applications. *Int. J. Electron. Electr. Eng.* **3**(2), 134–143 (2015)
9. Flattie, A.G.: Effect of high altitude aeronautical platforms with cognitive relay for radar performance. *Int. J. Sig. Process. Syst.* **3**(2), 159–165 (2015)
10. Motlagh, N.H., Taleb, T., Arouk, O.: Low-altitude unmanned aerial vehicles-based internet of things services: comprehensive survey and future perspectives. *IEEE Internet Things J.* **3**(6), 899–922 (2016)
11. Mehmood, Y., Gurg, C., Muehleisen, M., Timm-Giel, A.: Mobile M2M communication architectures, upcoming challenges, applications, and future Directions. *EURASIP J. Wirel. Commun. Netw.* (1), 1–37 (2015). <https://doi.org/10.1186/s13638-015-0479-y>
12. Afanasieva, L., Minochkin, D., Kravchuk, S.: Providing telecommunication services to antarctic stations. In: Proceedings of the 2017 International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo), Odessa, Ukraine, 11–15 September 2017, pp. 1–4. IEEE Conference Publications (2017). <https://doi.org/10.1109/UkrMiCo.2017.8095419>. (IEEE Xplore Digital Library)
13. Zgurovsky, M., Ilchenko, M., Kravchuk, S., Kotovskiy, V., Narytnik, T., Cybulskiy, L.: Prospects of using of aerial stratospheric telecommunication systems. In: Proceedings of the 2016 IEEE International Scientific Conference “RadioElectronics & InfoCommunications” (UkrMiCo 2016), Kyiv, Ukraine, 11–16 September 2016. IEEE Conference Publications (2016). <https://doi.org/10.1109/UkrMiCo.2016.7739636>. (IEEE Xplore Digital Library)
14. Kravchuk, S., Minochkin, D., Omiotek, Z., Bainazarov, U., Weryńska-Bieniasz, R., Iskakova, A.: Cloud-based mobility management in heterogeneous wireless networks. In: Proceedings of the Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments, SPIE, Wilga, Poland, 7 August 2017, vol. 10445, p. 104451W (2017). <http://dx.doi.org/10.1117/12.2280888>
15. Almási, B.: Multipath communication – a new basis for the future internet cognitive infocommunication. In: 2013 IEEE 4th International Conference on Cognitive Infocommunications (CogInfoCom), Budapest, Hungary, 2–5 December 2013, pp. 201–204 (2013). <https://doi.org/10.1109/CogInfoCom.2013.6719241>

16. Garg, A., Das, S.S.: Design of troposcatter broadband link based on SCFDE. In: 2017 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS), pp. 1–6 (2017). <https://doi.org/10.1109/ANTS.2017.8384095>
17. Li, C., Chen, X., Liu, J., Liu, Z.: One-way time transfer for large area through tropospheric scatter. In: 2017 17th IEEE International Conference on Communication Technology (ICCT), pp. 1–5 (2017). <https://doi.org/10.1109/ICCT.2017.8359472>
18. Li, C., Chen, X., Xie, Z.: A closed-form expression of coherence bandwidth for troposcatter links. *IEEE Commun. Lett.* **22**(3), 646–649 (2018). <https://doi.org/10.1109/LCOMM.2017.2785850>
19. Maokai, H., Xihong, C., Tao, S., Shaoqiang, D.: New generation troposcatter communication based on OFDM modulation. In: Ninth International Conference on Electronic Measurement & Instruments (ICEMI 2009), pp. 3:164–3:167 (2009)
20. Xie, Z., Chen, X., Liu, X., Zhao, Y.: MMSE-NP-RISIC-based channel equalization for MIMO-SC-FDE troposcatter communication systems. *Math. Prob. Eng.* 1–9 (2016). <http://dx.doi.org/10.1155/2016/5158406>. Hindawi Publishing Corporation, ID 5158406
21. Kravchuk, S., Kaidenko, M.: Features of creation of modem equipment for the new generation compact troposcatter stations. In: Proceedings of the International Scientific Conference “RadioElectronics & InfoCommunications” (UkrMiCo 2016), Kyiv, Ukraine, 11–16 September 2016, pp. 365–368. *IEEE Conference Publications* (2016). <https://doi.org/10.1109/UkrMiCo.2016.7739634>. (IEEE Xplore Digital Library)
22. Kravchuk, S.O.: Principles for creating portable tropospheric radio relay stations. In: Proceedings of the 9th International Scientific Conference Modern Challenges in Telecommunications, Kyiv, Ukraine, 21–25 April 2015, pp. 254–256 (2015). (in Russian)
23. Kravchuk, S.O., Kaidenko, M.M.: Modem equipment for the new generation compact troposcatter stations. *Inf. Telecommun. Sci.* **7**(1), 5–12 (2016)
24. Unkauf, F.: The next generation of troposcatter systems. *Raytheon Technol. Today* (3), 9–10 (2007)
25. Bastos, L., Wietgreffe, H.: Highly-deployable troposcatter systems in support of NATO expeditionary operations. In: IEEE Conference Proceedings on Military Communications Conference MILCOM, Baltimore, 7–10 November 2011
26. Li, L., Wu, Z.-S., Lin, L.-K., Zhang, R., Zhao, Z.-S.: Study on the prediction of troposcatter transmission loss. *IEEE Trans. Antennas Propag.* **64**(3), 1071–1078 (2016)
27. Yun, Z., Iskander, M.F.: Ray tracing for radio propagation modeling: principles and applications. *IEEE Access* **3**, 1089–1100 (2015)
28. Nafisi, V., Madzak, M., Böhm, J., Schuh, H., Ardalani, A.A.: Ray-traced tropospheric slant delays in VLBI analysis. *Vermessung Geoinformation* (2), 149–153 (2011)
29. Valtr, P., Pechac, P.: Tropospheric refraction modeling using ray-tracing and parabolic equation. *Radioengineering* **14**(4), 98–104 (2005)
30. Zhao, X., Yang, P.: A simple two-dimensional ray-tracing visual tool in the complex tropospheric environment. *Atmosphere* **8**(35), 1–10 (2017)
31. Dinc, E., Akan, O.B.: A ray-based channel modeling approach for MIMO troposcatter beyond-line-of-sight (b-LoS) communications. *IEEE Trans. Commun.* **63**(5), 1690–1699 (2015)
32. Dinc, E., Akan, O.B.: A nonuniform spatial rain attenuation model for troposcatter communication links. *IEEE Wirel. Commun. Lett.* **4**(4), 441–444 (2015)
33. Dinc, E., Akan, O.B.: Beyond-line-of-sight ducting channels: coherence bandwidth, coherence time and rain attenuation. *IEEE Commun. Lett.* **19**(12), 2274–2277 (2015)
34. Dinc, E., Akan, O.B.: Fading correlation analysis in MIMO-OFDM troposcatter communications: space, frequency, angle and space-frequency diversity. *IEEE Trans. Commun.* **63**(2), 476–486 (2015)

35. Barclay, L. (ed.): Propagation of Radiowaves, p. 470. The Institution of Engineering and Technology, London (2013)
36. Athanaileas, T.E., Athanasiadou, G.E., Tsoulos, G.V., Kaklamani, D.I.: Parallel radio-wave propagation modeling with image-based ray tracing techniques. *Parallel Comput.* **36**, 679–695 (2010)
37. Kaur, B.: MATLAB and K-wave based outdoor ray propagation predictor tool SNELLIX for surface wave modelling. *Innov. Syst. Des. Eng.* **6**(11), 7–18 (2015)
38. Ilchenko, M., Kravchuk, S., Minochkin, D., Afanasieva, L.: Troposcatter communication link model based on ray-tracing. *Inf. Telecommun. Sci.* (2), 15–20 (2018). <https://doi.org/10.20535/2411-2976.22018.15-20>
39. Ilchenko, M.E., Kravchuk, S.A.: Information telecommunication broadband radio access systems. *J. Autom. Inf. Sci.* **38**(4), 69–77 (2006). <https://doi.org/10.1615/JAutomatInfScien.v38.i4.60>



Prospects for the Development of Geostationary Satellite Communications Systems in the World

T. Narytnyk^(✉) and S. Kapshtyk^(✉)

National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Peremoga Avenue 37, Kyiv 03056, Ukraine
director@mitris.com, kapshtyk@gmail.com

Abstract. The current state and prospects of development of GEO satellite communication systems are considered in the context of the market of satellite communication services and used in the space segment technical solutions. The features of the development of geostationary satellite systems using High Throughput Satellites are considered. The features of the development of systems with High Throughput Satellites of various operators are shown. Special attention is paid to the use of high-throughput satellite technologies in Intelsat-EPIC satellites in the C and Ku frequency bands. The advantages of Intelsat EPIC solutions are shown. Trends in the development of Regular GSO Satellites are considered. The trends in combining High-Throughput Satellite and Regular Satellites in one orbital position are shown.

Keywords: Satellite communication line · Geostationary orbit · High-Throughput Satellite · Regular satellite

1 Features of the Regular Geostationary Satellites and the High-Throughput Satellites

At the turn of the XX–XXI centuries, due to advances in many fields of technology and technology, a typical view of the requirements, specifications and architecture of on-board equipment for the geostationary (GSO) civil/universal communications satellite was formed. The satellite has to be equipped with single/multy band payload in L, S, C, Ku, Ka-band (depending on the satellite service: FSS, RCC, MCC), provides services in several semi-global or regional beams formed by antennas with a contour beams. In each beam, the service is provided by several transponders, which is achieved through two or three times use of the allocated frequency band. The total bandwidth of a regular satellite does not exceed 10 GHz. Depending on the frequency band of the satellite, the requirements for the accuracy of the kipping of the orbital position ($\pm 0,1^\circ$ for the C band and $\pm 0,05^\circ$ for Ku band) are specified. The planned lifetime of GSO satellites is not less 15 years.

DVB-S satellite communications (including DVB-S2, DVBV-S2x) standards performed by the DVB standardization committee [2] allowed to unify the requirements

for RF-equipment of the GEO satellite payload and to form a typical model of the satellite communication line [1].

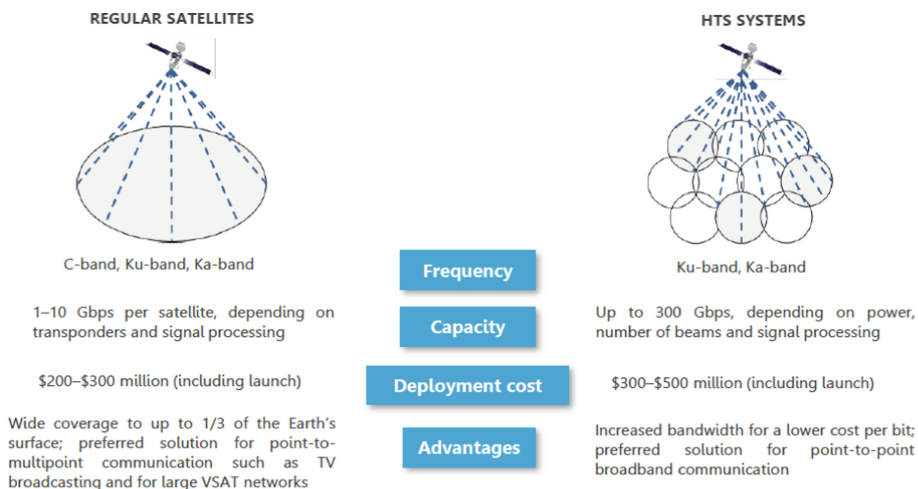


Fig. 1. Concepts of the FSS Regular Satellite and High Throughput Satellites systems (HTS Systems)

The Regular Satellite (Fig. 1) has a wide service area, uses frequency bands C (4/6 GHz), Ku (11/14, 12/17 GHz), Ka (20/30 GHz) (one or several simultaneously). The total bandwidth of a Regular Satellite is 1 ÷ 10 Gbps. The deployment cost (including launch and launch insurance) is estimated at \$200–300 million. The advantage of these satellites is the ability to effectively transmit information from one correspondent to many correspondents (point-multipoint, broadcasting mode). The classic branch of the use of FSS and BSS Regular Satellites is satellite broadcasting and the providing of other types of services for the distribution of video information.

The High-Throughput Satellite (HTS satellite) forms a service area with a number of spot beams with beamwidth typically 1 ÷ 2 or less. HTS satellites use Ku (11/14 GHz) and Ka (20/30 GHz) frequency bands with multitime frequency reuse of the dedicated frequency band. Euroconsult distinguishes two HTS satellites generations: first-generation satellites with capacity up to 100 Gbps, and second-generation satellites with capacity up to 300 Gbps and more [1]. The cost of deployment (including launch and launch insurance) of HTS satellite in the geostationary orbit is \$300–500 million. The advantages of satellite HTS include the low cost of information transmission per bits of transmitted information, which determines their high efficiency in organizing the transfer of information between two separate correspondents (point-to-point mode “Point-to-point”). But the available for one subscriber (end-user) bandwidth does not exceed 5% of the total satellite HTS bandwidth.

Another difference in the system with HTS satellites is the presence of special beams to Gateway (Gateway Beams) or teleports that are intended for switching HTS

system space segment to Ground Backbone and routing traffic. These beams are directed on the teleport location area/areas. Unlike HTS Satellites, in

Telecommunication Systems with Regular Satellites, the Teleport or the Hub-Station is usually in the same beam with user stations.

During the period from January 1, 2016 to July 5, 2017, 42 commercial communications satellites were launched on the GEO (see Table 1). Of these, 8 satellites belong to the class of HTS satellites: Intelsat-29e, Intelsat-32e, Intelsat-33e, Intelsat-35e, EchoStar-19, Inmarsat-5 F4, ViaSat-2, Eutelsat-172B (see note in Table 1). The rest are Regular Satellites.

Table 1. List of commercial communications satellites successfully launched into the geostationary orbit from January 1, 2016 to July 5, 2017

Date	The name of the satellite	Date	The name of the satellite
15.01.2016	Belintersat-1	5.01.2017	TJS 2
27.01.2016	Intelsat 29e (*)	28.01.2017	Hispasat AG1
29.01.2016	Eutelsat 9B	14.02.2017	Intelsat 32e/SkyBrasil-1 (*), Telkom-3S
4.02.2016	SES-9	16.03.2017	Echostar 23
9.03.2016	Eutelsat 65 West A	30.03.2017	SES-10
6.05.2016	JCSAT-14	12.04.2017	Shijian 13
27.05.2016	Thaikom 8	4.05.2017	Koreasat-7, SGDC-1
9.06.2016	Intelsat 31, DLA-2	5.05.2017	GSAT-9
15.06.2016	Eutelsat 117 West B, ABS-2A	15.05.2017	Inmarsat-5 F4 (*)
18.06.2016	EchoStar 18, BRISat	18.05.2017	SES-15
5.08.2016	Tiantong-1 01	1.06.2017	ViaSat-2 (*), Eutelsat 172 B (*)
14.08.2016	JCSAT-16	5.06.2017	GSAT-19
24.08.2016	Intelsat 33e (*), Intelsat 36	8.06.2017	Echostar 21
5.10.2016	NBN-Co/Sky Muster II, GSAT-18	23.06.2017	BulgariaSat-1
22.11.2016	Tianlian I-04	28.06.2017	EuropaSat/HellasSat-3, GSAT-17
18.12.2016	EchoStar 19 (*)	5.07.2017	Intelsat 35e (*)
21.12.2016	Star One D1, JCSAT-15		

Note: (*) marked satellites that relate to the class of HTS satellites.

2 Features of Modern Geostationary High-Throughput Systems

During the one and a half year period, 8 HTS Satellites were launched on the GEO. Of these, the original definition of a “High-Throughput Satellite” [1] fully corresponds to three satellites: EchoStar-19, Inmarsat-5 F4, ViaSat-2.

2.1 EchoStar System

The EchoStar-19/Jupiter 2 Satellite (operator: EchoStar Satellite Services LLC, manufactured by Space System Loral) is built in base of the SSL-1300 bus and launched on the GEO using the Atlas-V launch vehicle on December 18, 2016). The Satellite is placed in the orbital position 97.1 W. The Satellite is designed to provide broadband services in the Ka-band (20/30 GHz). At the launch date, the EchoStar-19 Satellite was the largest satellite in terms of the total capacity and the number of spot beams. The total capacity of the Satellite is 220 Gbps, which is almost twice the capacity of the previous EchoStar-17/Jupiter 1 Satellite with capacity 120 Gbps. But this is not the limit. According to EchoStar announcement, the EchoStar-24/Jupiter 3 Satellite, which is also under construction by Space System Loral, will provide capacity of 500 Gbps [3]. In the long term, according to Euroconsult forecasts, there are expected satellites with capacity up to 1 Tbps [4]. The EchoStar-19/Jupiter 2 satellite service area is formed with 130 spot beams and covers the entire territory of the United States (see Fig. 2a).



Fig. 2. (a) EchoStar-19/Jupiter 2 satellite service area; (b) ground infrastructure for managing and providing broadband internet access services

EchoStar-19/Jupiter 2 is based on the JUPITER™ technology developed by Hughes Network Systems, LLC. and implements HughesNet®Gen5's fifth generation broadband satellite access technology. JUPITER™ technology [3] is based on the new standard DVB-S2x and provides in the Forward Channel (Down-link) the Symbol Rate of Group Transport Stream up to 225 Msbps, which allows reaching the Transmission Rate of up to 450 Mbps, while in the Return Channel (Up-link) the Symbol Rate is up to 8 Msbps. The Adaptive Coding Modulation (ACM) mode is used with FM-4, FM-8, AFM-16 and AFM-32 modulation, in with AFM-64 modulation in future, with 5% roll-off factor in combination with Low Density Parity Check (LDPC) Forward Error Coding (FEC) with Code Rate up to 9/10. FM-4, FM-8 modulations and LDPC codes with Code Rate up to 9/10 are used in the Return Channel.

Operation of the EchoStar-19/Jupiter 2 Satellite and the providing of services are provided by the Ground Infrastructure of EchoStar Satellite Services L.L.C, which consists of two Control Centers and four Teleports (see Fig. 2b).

2.2 ViaSat System

ViaSat-2 Satellite (operator: Viasat Inc., manufacturer: Boeing Satellite Systems) is based on the BSS-702HP bus and launched into orbit using the Arian-V launch vehicle (Arianespace launch operator, launch date July 1, 2017). The Satellite was placed in the orbital position of 69.9 W. Today, the ViaSat-2 Satellite remains the most powerful satellite which was produced by Boeing Satellite Systems in base of the BSS-702HP platform. The Satellite is designed to provide Broadband Internet Access in Ka-band (20/30 GHz). The total capacity of the Satellite is about 300 Gbps, which is two and a half times the capacity of the previous ViaSat-1 Satellite (ViaSat-1 Satellite Capacity is 130 Gbps). Given the prospects for increasing the demand for Satellite Capacity in the North American region, with the goal of providing global coverage, ViaSat Inc. announced plans to achieve the Capacity of 1Tbps on each of the three satellites of ViaSat-3 series [9].

The ViaSat-2 Satellite provides Broadband Internet Access in Ka-band over North America with 72 spot beams (B-Type Spot Beam). The Satellite Service Area has grown more than seven times [7]. To connect Space Segment to the Terrestrial Network through the GateWay station, the Satellite has 20 GateWay Beams (A-Type Spot Beam) (ViaSat-2 satellite service area is shown in Fig. 3).

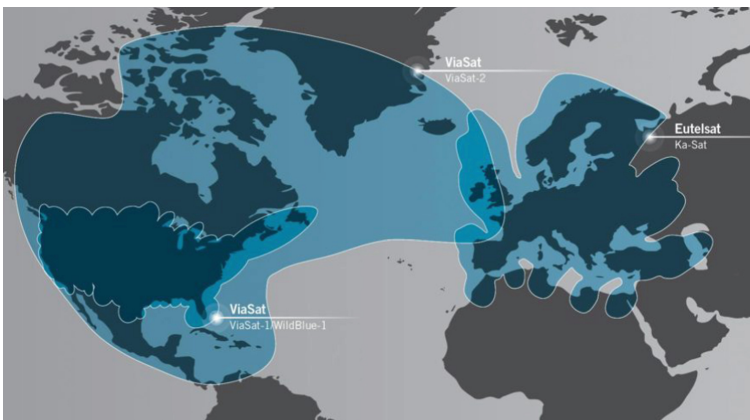


Fig. 3. The ViaSat-2 satellite service area and the broadband access service area in the base of ViaSat broadband access technology

The Satellite uses SurfBeam®2 Technology to provide a wide range of information services for consumers in the base of broadband Internet access. The satellite’s payload is constructed according to the “bent-pipe” scheme with asymmetric lines in the forward and return directions. In the direction to the end-user (Forward line) in the

500 MHz frequency bandwidth, the total information stream is transmitted at a single frequency in the time-division mode (TDMA) mode with Symbol Rate up to 416.67 Msbps. The set of frequencies with time compression at each frequency is used in the return channel. The system uses ACM. In the Forward Channel, the following types of modulation are used: AFM-16, FM-8, FM-4, in the Return Channel: FM-8, FM-4, FM-2.

The advantages of SurfBeam@2 technology are the Capacity of the System based on the System Service Class, rather than the User Terminal. The on-line services, such as the transmission of video and voice information in real time, have privilege. The maximum Information Rate for information transmission to one User Terminal in the direction from the Hub-station (Forward Channel) is up to 12 Mbps, in the opposite direction to 3 Mbps [8]. ViaSat Inc. aims to provide a global coverage of the surface of the Earth not only in Ku-band (11/14 GHz), but also in the Ka-band (20/30 GHz).

There is a new trend for operators of HTS satellites the combination of service areas of different satellites to expand the joint service area based on the shared SurfBeam@2 DOCSIS® information technology for User Terminals. Figure 3 shows shared service area in the Ka-band of the ViaSat-2 and Ka-Sat satellites.

2.3 Inmarsat System

Inmarsat-5 F4 satellite (operator: Inmarsat plc., UK, manufacturer: Boeing Satellite Systems) is based on the BSS-702HP bus [11] and launched into orbit using the Falcon-9 launch vehicle (Launch service operator: SpaceX, launch date May 15, 2017). The Inmarsat-5 F4 is the fourth satellite in a series of four HTS satellites that create the Space Segment of the Global Xpress system. The main payload of the Inmarsat-5 F4 Satellite in the Ka-band consists of 89 transparent (bent-pipe) transponders. The features of the Inmarsat-5 F4 satellite are the accommodation of Hosted Payload, which operates in two L- and Ka-bands, and the Customer of which is the US Government. The Inmarsat-5 F4 Satellite was considered as a Standby Satellite in the Geostationary orbit at launch time, according to official statements by Inmarsat plc. The satellite is temporarily placed in the orbital position of 83.4°E.

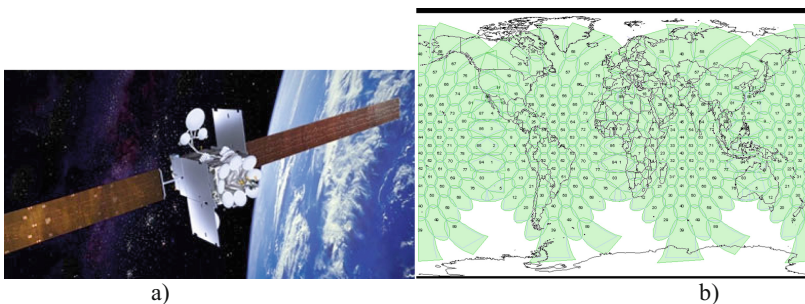


Fig. 4. (a) The Inmarsat-5 F4 satellite: the art image of the satellite in operation mode; (b) The Inmarsat-5 F4 satellite: the coverage area of Inmarsat-5 satellites in the Global Xpress system

The Inmarsat-5 F4 Satellite forms 89 fixed and super-narrow beams of circular and elliptical shape (Fig. 4b) and eight steerable super-narrow beams, two of which are designed to provide communication with GateWay stations in the Ka-band.

The Satellite is intended to provide high-speed data transmission and Internet access services in Global Xpress System for mobile consumers on land, at sea and in the air. Each Fixed Beam provides a total Data Transmission Stream with Bit Rate 100 Mbps, which is shared between the down-link and up-link. Each User Terminal within a single beam provides information receiving with Bit Rate up to 50 Mbps and information transmission with Bit Rate up to 5 Mbps. Steerable beams are capable to provide information transmitting at Bit Rate of several hundred Mbps in channels with a bandwidth of 40, 125, 270, 400 and 730 MHz for military and commercial consumers (for commercial users up to 600 Mbps). Figure 5 shows the scheme of providing services through the Steerable Beams of Inmarsat-5 Satellites.

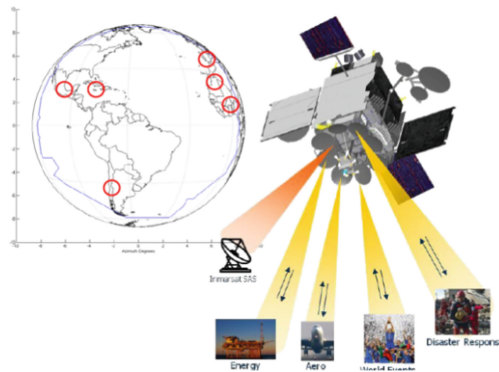


Fig. 5. Organization of Inmarsat-5 satellite services with a steerable beam in Ka-band

The Global Xpress System provides iDirect-based Broadband Access Services and uses a Ground Infrastructure that includes 6 Teleports located within the service areas (Fig. 6a). In the Forward Channel (to the end-user terminal), the iDirect protocol, which uses the DVB-S2 standard in the single-frequency timed-mode mode, is used, the multi-carrier MF-TDMA mode with TDMA individually for each carrier is used in the Return Channel. Teleports are connected by a ring of terrestrial high-speed fiberoptic lines, which ensures high reliability of the system and the high quality of service provision (see Fig. 6b).

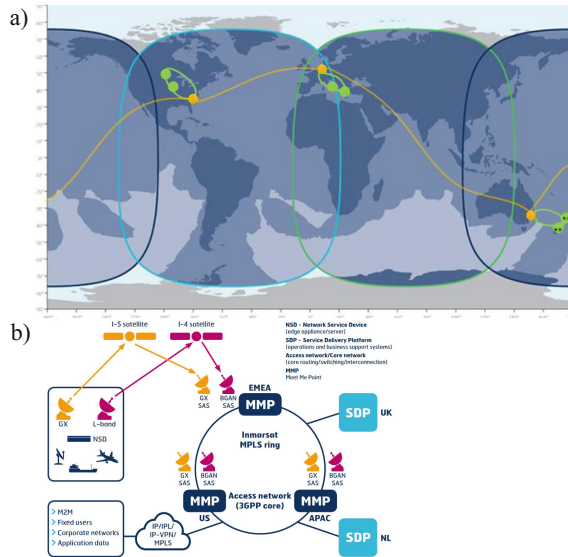


Fig. 6. Inmarsat system ground infrastructure to provide Global Xpress services: (a) Teleports and ground fiber optic lines; (b) service delivery scheme

2.4 Intelsat System

The world's leader in satellite communications, Intelsat S.A., launched the new GEO Satellite System Intelsat EPIC^{NG} during 2016–2017. The Space Segment of the System consists of four Intelsat EPIC^{NG} series satellites: IS-29e (launched January 27, 2016, orbital position 50°W), IS-33e (launched August 24, 2016, orbital position 60°E), IS-32e (launched February 14, 2017, orbital position 43.1°W), and IS-35e (launched July 5, 2017, orbital position 34.5°W). The IS-29e, IS-33e and IS-35e of the Intelsat EPIC^{NG} series are based on the BSS-702MP (manufacturer: Boeing Satellite Systems). The IS-32e satellite is based on the Eirostar-3000X bus (manufacturer of Airbus Defense and Space).

Intelsat EPIC^{NG} Satellites have a number of differences from HTS satellites of other operators such as EchoStar, ViaSat, Inmarsat. Intelsat EPIC^{NG} Satellites use the technology to create a service area by the large number of spot beams, characteristic for satellites in more high bands in the Ka-band, in lower frequency bands C (4/6 GHz) and Ku (11/14 GHz). The Intelsat EPIC^{NG} Satellites have a combined payload that uses all frequency bands allocated for fixed satellite service: C, Ku, Ka (see Table 2).

Table 2. Parameters of the payload of Intelsat EPIC^{NG} series satellites

Satellite	IS-29e	IS-32e	IS-33e	IS-35e
Orbital position	50°з.д.	43,17°з.д.	60°с.д.	34,5°з.д.
Range C	Frequency band (MHz)	864		2670
	Polarization	Linear, vertical/horizontal		Linear, vertical/horizontal circular, right/left
	EIRP (dBW)			
	Spot beam			46,2 ÷ 52,4 41,0 ÷ 43,5
	Wide beam	>36,0		33,3 ÷ 37,5
	G/T (dB/K)			
	Spot beam			2,6 ÷ 12,8 -1,6 ÷ 1,5
	Wide beam	2,7		-10,3 - 7,2
	Number of transponders	14		20
	Ku band	Frequency bandwidth (MHz)	9395	
Polarization		Linear, vertical/horizontal	Linear, vertical/horizontal*	Linear, vertical/horizontal
EIRP (dBW)				
Spot beam		46,7 ÷ 63,4	44,0 ÷ 61,5	48,7 ÷ 61,6
Wide beam		47 ÷ 48,2	45,7 ÷ 48,7*	43,6 ÷ 45,3
G/T (dB/K)				
Spot beam		7,1 ÷ 17,2	6,5 ÷ 15,8	7,0 ÷ 17,0
Wide beam		-1,0 ÷ 2,0	-1,4 ÷ 3,1	-3,3 ÷ -0,7
Number of transponders		56	60 (20*)	
Ka band		Frequency bandwidth (MHz)	450	
	Polarization	circular, right/left		circular, right side/left side
	EIRP (dBW)	>35,5		>35,3
	G/T (dB/K)	Up to 6,6		Up to -7,0
	Number of transponders	1	21	1

Note: * for the IS-32e satellite, the EIRP and G/T are given only for the Intelsat EPIC^{NG} payload.

The Intelsat S.A. creates **Intelsat EPIC^{NG} System** as a new innovative platform to provide a new satellite services and to create a different architecture networks in C-, Ku- and Ka-bands [15, 16]. The service area of the Intelsat EPIC^{NG} System covers land and main aviation and maritime routes (Fig. 7). Such a form of the Service Area is due to the new segment of satellite services: broadband access for airplanes (in-flight connection) on the main routes of transoceanic flights, and seagoing vessels on the high seas and international routes.

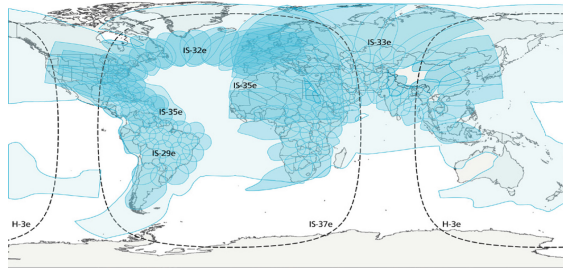
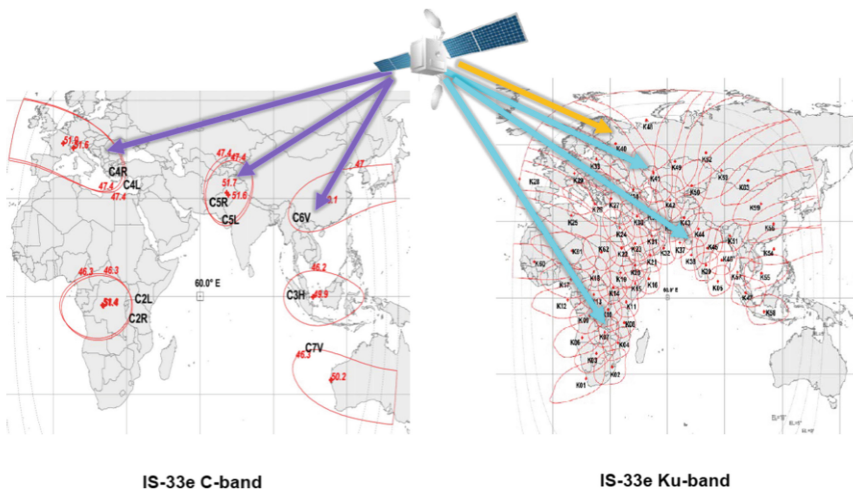


Fig. 7. Intelsat Epic^{NG} system service area

The feature of the Intelsat EPIC^{NG} System is the ability to create a communications network of any configuration, including the “Star” and “Mash” topology networks with “one hop” connection. Networks may include earth stations that are located within the same beam, stations located in other beams, stations operating in the same or in different bands. Figure 8 shows the principle of networks organizing by the Intelsat-33e Satellite in two frequency bands: C and Ku.



IS-33e C-band

IS-33e Ku-band

Fig. 8. Principle of Intelsat EPIC^{NG} system organization

To ensure the provision of connections in “one hop” mode, the Intelsat EPIC^{NG} Payload is equipped with channel forming and switching equipment based on the Transparent Digital Switch Matrix (Fig. 9a). The switching matrix performs switching of subchannels with bandwidth of 2.6 MHz each in the frequency band 1.4 ... 1.9 GHz (L-band). The bandwidth of 500 MHz is broken down into 192 subchannels. Depending on the desired bandwidth, which is determined by the width of the dial-up spectrum spectrum, one or more adjacent subchannels are selected. The following types of connections are used (Fig. 9b): Route - transfer of the frequency band of one or more adjacent subchannels from the “ground-to-space” line (Up-Link) to the “space-to-ground” line (Down Link), Fan-Out - duplication of the sub-channel frequency band or adjacent subchannels the “ground-to-space” line (Up-Link) in two frequency bands of the “space-to-ground” line (Down Link), Fan-In - combining of two spaced-apart subchannels or groups of neighboring sub-channels from “ground-to-space” line in one band of frequencies in the “space-to-ground” line, Order Wire - simultaneous transmission of not crush subchannel pools or groups of sub-channels of the “ground-to-space” line in two spaced strips of the “space-to-ground” line.

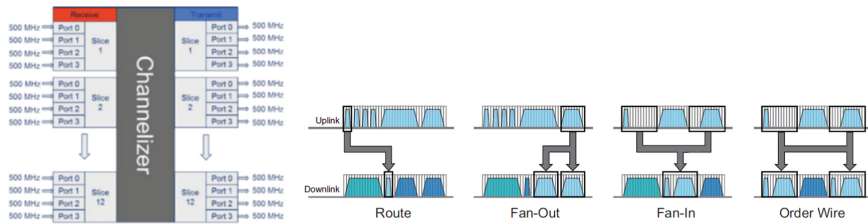


Fig. 9. The principle of the channels formation and switching of in the Intelsat EPIC^{NG} payload: (a) the channels formation and switching; (b) the types of switching

The fundamental difference of the **IS-29e** and **IS-33e** Satellites is the application of HTS Satellite Service Area forming technology in the Ku-band. In particular, the IS-33e Satellite has 63 fixed circular or elliptical shape spot beams (Fig. 10a). Additionally, individual rays are allocated to the gateway stations. In addition, the IS-29e and IS-33e have narrow beams in the C-band and a global beam in the Ka-band (see Fig. 10b, c).

The features of the **IS-35e** Satellite (orbital position 34°E) is the application of HTS Satellite technology in the C-band. The Satellite has wide beam and 17 spot beams in this frequency band (Fig. 11a). The Satellite has three broad regional beams in Ku-band (Fig. 11b, c, d).

The features of the **IS-32e** satellite include not only the utilization of the European manufacturer’s bus for satellite construction, but also the joint ownership of the Ku band resource and the hosted payload of the Ka band. This is reflected in the satellite name: Intelsat-32e/SKY Brasil-1. The 60 Ku-band transponders are designed to provide direct satellite broadcasting services in Brazil in one wide and 20 spot beams. At the same time, some of them can be used to ensure the workload of Intelsat EPIC^{NG} in

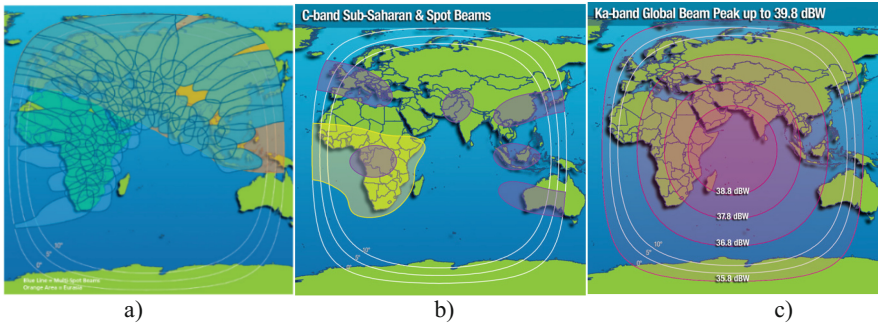


Fig. 10. S-33e Satellite Service Areas: (a) 63 spot user's beams in Ku-band and 7 GateWay beams; (b) 6 spot beams and 1 regional beam in C-band; (c) one global beam in Ka-band

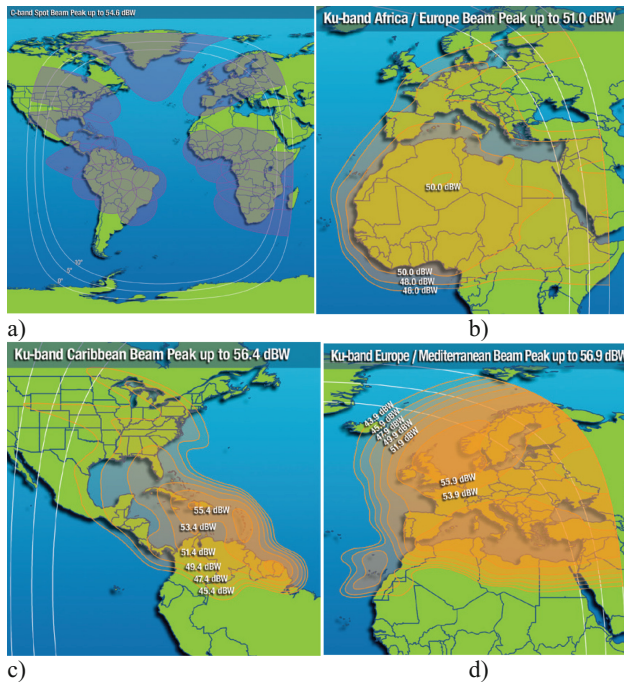


Fig. 11. IS-35e satellite payload service areas: (a) C-band; (b–d) Ku-band

16 spot beams [13, 14]. Due to this satellite payload Intelsat S.A. company increases satellite capacity to provide broadband satellite access in the region of maritime and aviation routes in the North Atlantic region (Fig. 12). Co-owners of the Ku-band payload are Intelsat S.A. and DirectTV Latin America – the Latin American division of AT&T/DirectTV. The owner of the Ka-band payload is YahSat, United Arab Emirates.



Fig. 12. Intelsat EPIC^{NG} IS-32e satellite payload service area

2.5 Eutelsat System

The features of satellites that were launched during the period considered is the combined payload, which includes the typical of HTS Satellite Payload and the typical of Regular Satellite Payload. **Eutelsat 172 B** (operator: Eutelsat SA, manufacturer of Airbus Defense and Space) launched on June 1, 2017 with the aid of the Arian-5 launch vehicle (Launch Operator: Arianespace) and placed in the orbital position of 172°E. The satellite is manufactured in base of Eurostar-3000EOR bus and equipped with dual-band payload (Fig. 13a). The Eutelsat 172B will provide services using 11 elliptical beams to provide airplanes with high-speed broadband access services in Ku-band which is based in the HTS Satellite Technology (Fig. 13g). In addition to the HTS payload the Satellite is equipped with the Regular Payload, which operates in C- and Ku-bands (Fig. 13b, c, d, e, f, h).

The **Eutelsat 65 West A** Satellite (orbital position 65°W) (based in the SSL-1300 bus) (Figs. 13, 14a) is three-band combination payload satellite. The satellite is equipped with 24 transponders in Ka-band with spot beams of elliptical shape to provide widthband access satellite services in South America (Fig. 14b). In addition, the satellite is equipped with the C-band (10 transponders with a bandwidth of 54 MHz each) and Ku-band (24 transponders with a bandwidth of 34 MHz each) payload of the regular configuration (with transponders of “transparent” type and antenna with shaped beam (Fig. 14b, c, d, e).

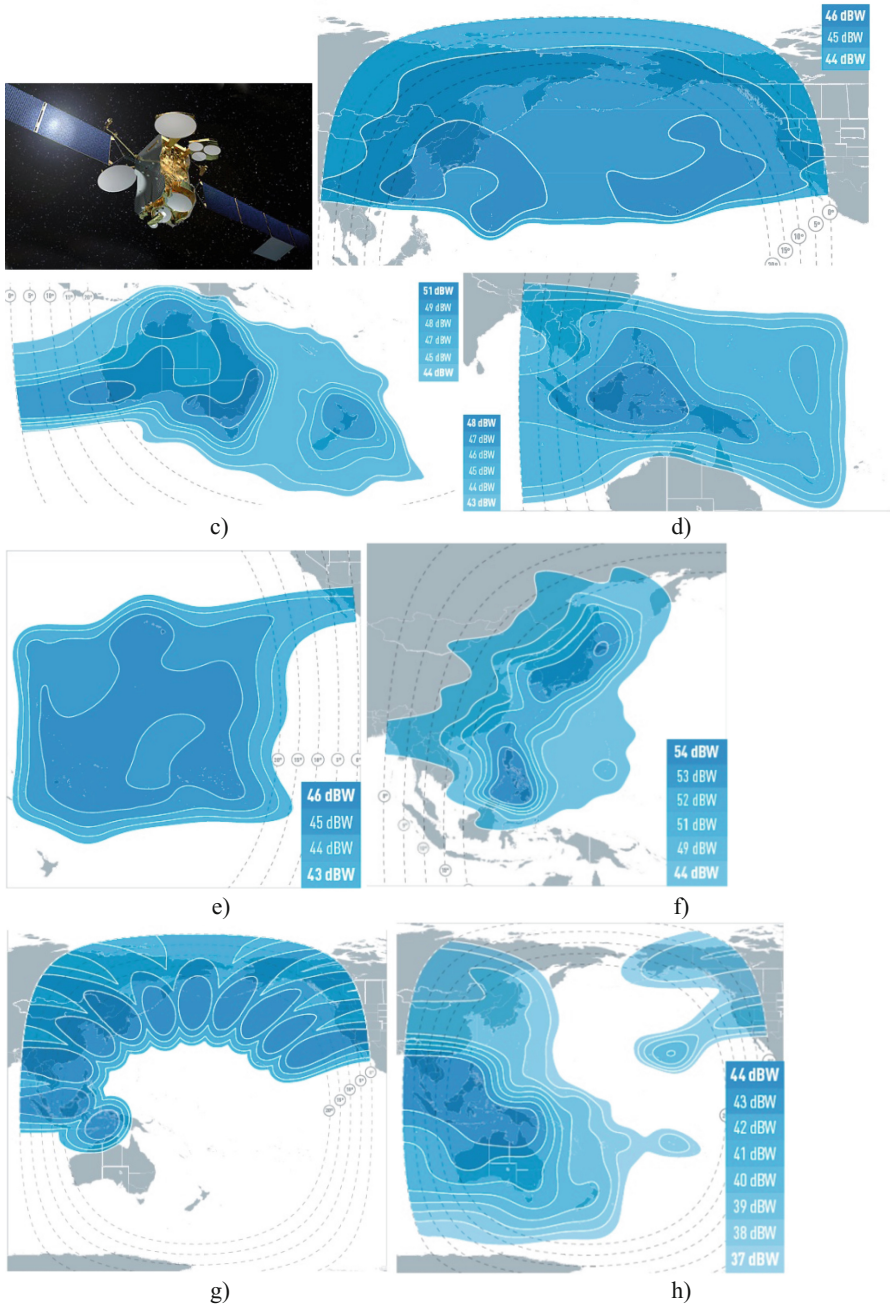


Fig. 13. The Eutelsat 172 B satellite: (a) the satellite's art picture; (b) the Ku-band "North Pacific" beam; (c) the Ku-band "South Pacific" beam; (d) the Ku-band "Southwest Pacific" beam; (e) the Ku-band "South-Eastern Pacific" beam; (f) the Ku-band "North-East Asia" beam; (g) the Ku-band HTS beams; (h) the C-band "Pacific Ocean" beam

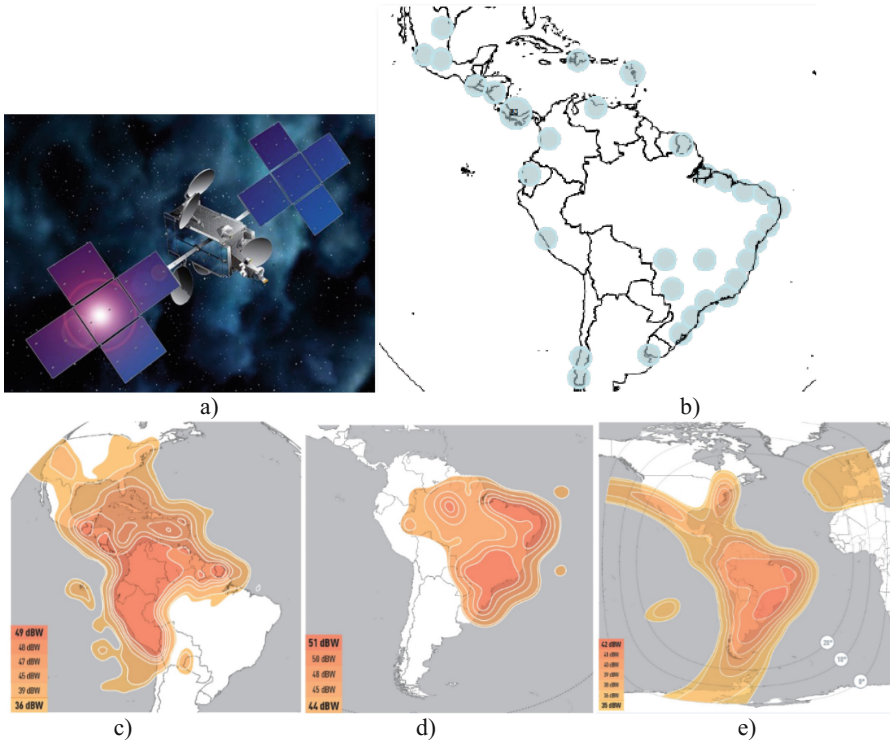


Fig. 14. The Eutelsat 65 West A satellite: (a) the satellite art picture; (b) Ka-band spot beams; (c) Ku-band American beam; (d) Ku-band Brazilian beam; (e) C-band beam

3 Geostationary Systems of the Regular Configuration

Due to the accumulated C-and Ku-band frequency orbital resource in the GEO, the leading satellite operators continue to launch the regular satellites.

3.1 Intelsat

Intelsat has continued to launch new satellites with a regular payload, which uses multiple shaped beams and hemispheric beam and uses in Ku and C frequency band. This class includes Intelsat 31/DLA-2 Satellites (orbital position 95°W) which has one C-band beam and 12 Ku-band beams, Intelsat 36 (orbital position 68.5°E), which has shaped beams in C- and Ku-frequency bands.

The **Intelsat-36** Satellite (orbital position 68.5°E) (based in SSL-1300 bus, Fig. 15a) refers to a Regular Satellites with a payload of traditional configuration and is intended to provide Fixed Satellite Services in the region of Central and South Africa in two frequency bands: in the C-band the Satellite has 10 Transponders, and in the Ku-band has 34 Transponders with a level of EIRP at the center of the service area of 54.7 dBW (the service areas in the C-and Ku-bands are shown in Fig. 15b, c).

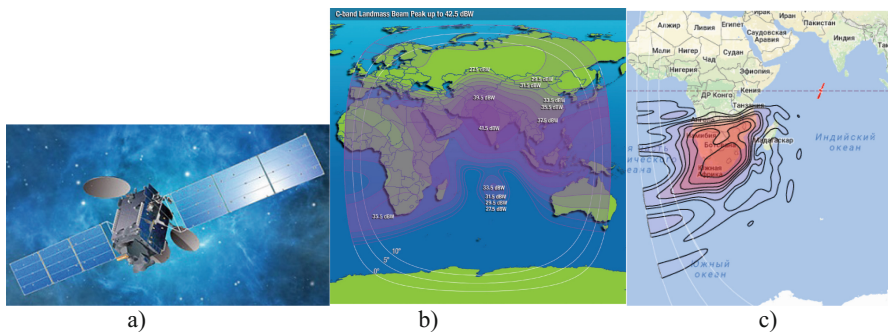


Fig. 15. Intelsat-36 satellite: (a) art picture; (b) C-band service area; (c) Ku-band service area

3.2 SES

SES S.A. (Luxembourg) continued to update its orbital group by launching satellites with regular configuration of payload in the Ku-band. The *SES-9 Satellite* (orbital position 108.2°E) (based on Boeing-702HP bus) is equipped with 81 Ku-band transponders and provides services in 7 shaped beams (Fig. 16).

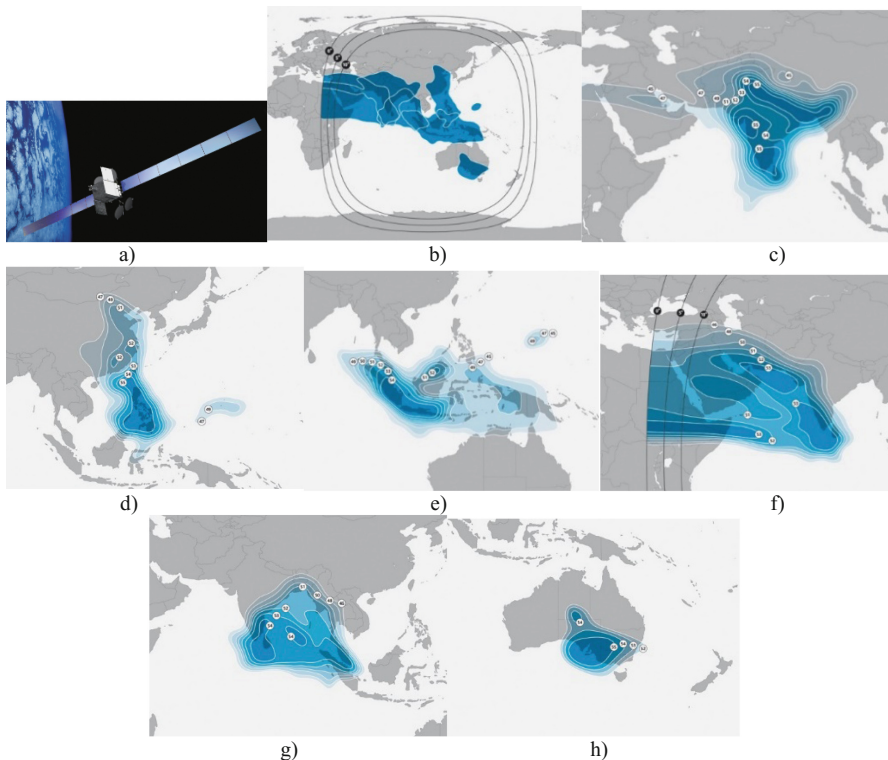


Fig. 16. Satellite SES-9: (a) art picture of the satellite in working mode; (b) Ku-band coverage area; (c) Ku-band “South Asia” beam; (d) Ku-band north-east beam; (e) Ku-band south-eastern beam; (f) Ku-band “Western Indian Ocean” beam; (g) Ku-band “Eastern Indian Ocean” beam; (h) Ku-band Australian beam

The *SES-10 Satellite* (orbital position 67°W) (based on Eurostar-3000 bus) (Fig. 17) is equipped with 55 Ku-band transponders and provides services in 4 shaped beams. The *SES-15 Satellite* (orbital position 129°W) is equipped with 16 Ku-band transponders and provides services in 2 shaped beams. Thus, SES focuses on the use of regular satellite technology in the Ku-band.

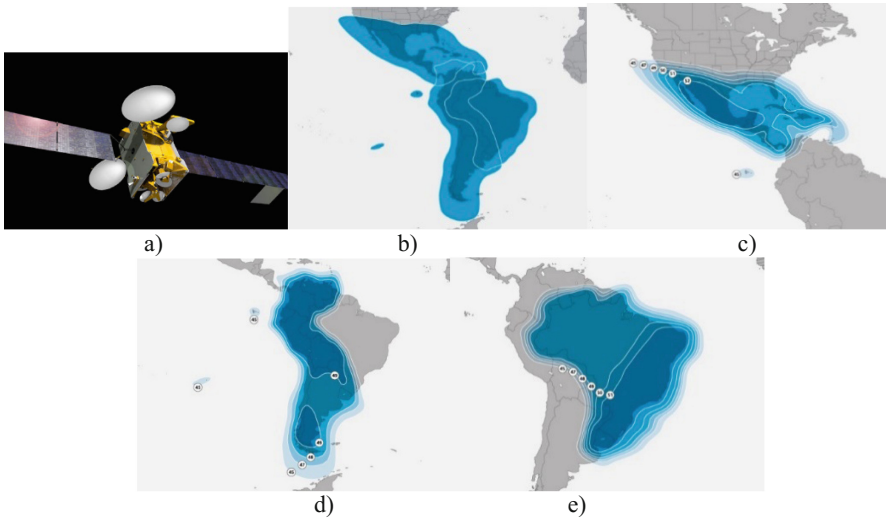


Fig. 17. The SES-10 satellite: (a) the art picture satellite in operation mode; (b) Ku-band beam; (c) Ku-band “Central America and Caribbean” beam; (d) Ku-band “South America” beam; (e) Ku-band Brazilian beam

3.3 Eutelsat

Eutelsat Company launched four satellites during 2016–2017. All satellites are designed to provide services in the Ku-band using shaped beam antennas. In particular, the **Eutelsat 9B** (orbital position 9°E) (based on Eurostar-3000 bus) has 50 Ku-band transponders and creates 6 beams (Fig. 18). The feature of providing services using the Eutelsat-9B satellite capacity is the combination in orbital one slot of regular powerful Ku-band Eutelsat-9B satellite which is designed to provide broadcast services in Ku-band and Ka-band powerful HTS Satellite KA-SAT that is designed to provide two-way satellite broadband access services in the Ka-band (Fig. 19). Additionally, the Eutelsat-9B is equipped with a hosted payload of the EDRS system - the European Data Relay System.

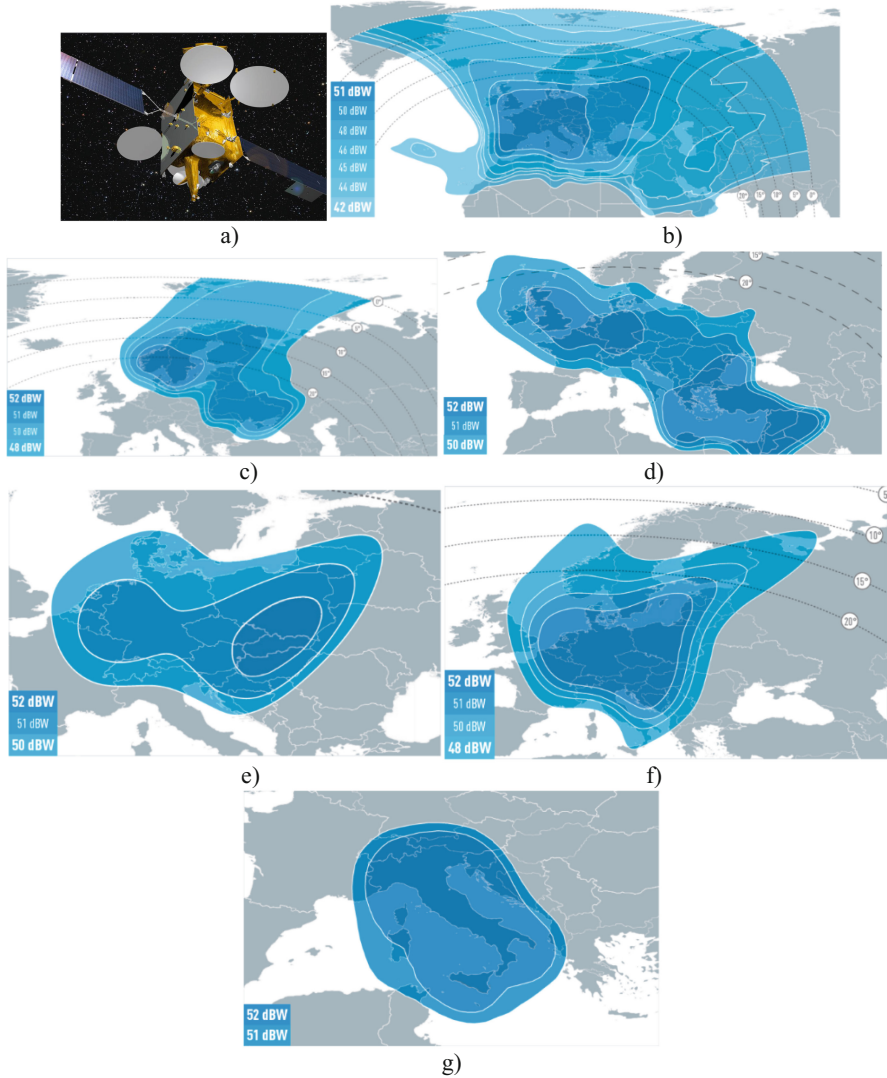


Fig. 18. Eutelsat-9B satellite: (a) art picture of satellite in operation mode; (b) Ku-band wide beam; (c) Ku-band North and Western Europe beam; (d) Ku-band Greece Extended beam; (e) Ku-band Germany A beam; (f) Ku-band Germany Ge beam; (g) Ku-band Italy beam

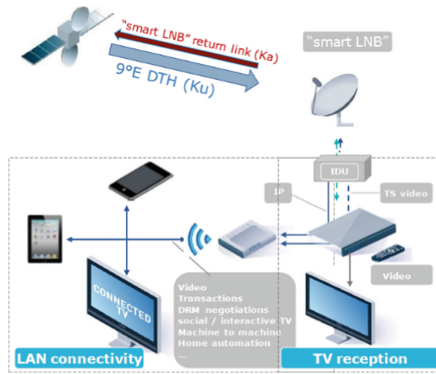


Fig. 19. The combination of Eutelsat-9B and KA-SAT satellites in the one orbital slot 9°E to provide a broadband and broadcast satellite service package

Eutelsat 117 West B Satellite (orbital position of 117°W) is manufactured on the base of Hughes - 702SP Bus (Full-electric Satellite Bus). The feature of this Bus is abandonment from chemical engines for attitude and orbit control and utilization for this mete of XIP propulsion system (Ion - Xenon Propulsion System - XIP). The Eutelsat 117 West B Satellite (Fig. 20a) behaves to the companions of the Fixed Satellite Service. Satellite Payload includes 48 transponders in Ku-band and service area is formed by 4 shaped antenna beams (Fig. 20b, c, d, e).

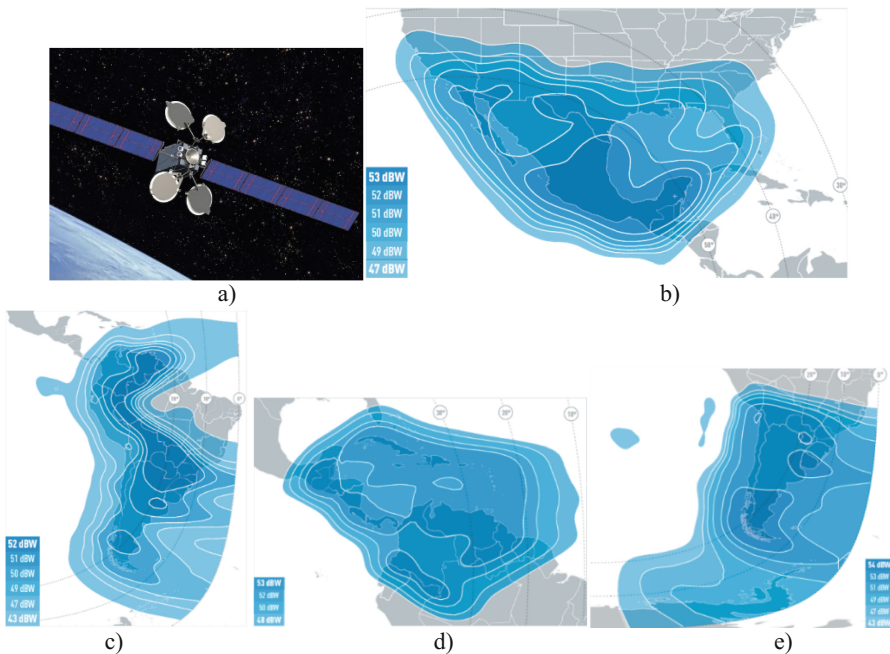


Fig. 20. The Eutelsat 117 West B satellite: (a) the art picture of satellite in operation mode; (b) Ku-band the beam 1; (c) the Ku-band beam 3; (d) Ku-band beam 5; (e) Ku-band beam

4 Conclusions

The analysis of technical performances and features of new telecommunication satellites launched on the GEO allows us to make the following conclusions:




1. Geostationary orbit remains the subject of special attention of satellite telecommunication systems operators. Telecom operators support and protect their own frequency resource on the GEO by updating their own satellite fleet.
2. An important direction in the development of the GEO orbital fleet of telecommunication satellites was the launch of High-Throughput Satellites, the capacity of the largest of them is approaching 1 Tbps. HTS satellite capacity grows in the Ka-band the most actively.
3. To accelerate the formation of a wider service area, operators of HTS satellites create common service areas by combining their own GSO satellites and implementing services based on common or compatible standards and broadband satellite communications technologies.
4. Taking into account the available frequency resource in the C-and Ku-bands on the GEO, the leading satellite telecommunication systems operators are introducing High Throughput satellite technologies, characteristic for the Ka-band, into new satellites in these frequency bands. The combination of technologies for the formation of a service area by high-energy spot beams and on-board inter-beam connectivity by switching the frequency channels and subchannels can significantly improve the efficiency of the satellite telecommunication system as a whole.
5. For the further development of high-speed/broadband satellite communications, interest is the using of higher frequency bands, in particular the V-band and terahertz frequency band and utilization of spread spectrum and ultra-spread spectrum signals. Given the peculiarities of the propagation of radio waves, these band can be applied not only for communication in the “ground-space” and “space-ground” lines, but also to ensure communication between satellites on the GEO and between satellites in lower orbits with satellites on GEO [15–17].
6. GEO Satellites with a Payload of the Regular configuration continue to occupy the dominant position in the GEO Communication Satellites Launch Manifests. The main purpose of these satellites is the replacement of satellites on the GEO, the term of operation of which is coming to an end. There are new trends in the development of Regular Satellites:
 - combined payload, which includes part of the Payload built on the HTS technology and the Regular Payload;
 - placing in one orbital position on the GEO at the same time a HTS Satellite and a Regular Satellite that offers for end user a combined package of broadband two-way access/communication service and multimedia/television broadcasting service.

References

1. Satellite Communications and Broadcasting Market Survey, 23rd edn, September 2016. © Euroconsult (2016)
2. Digital Video Broadcasting (DVB). Implementation guidelines for second generation system for Broadcasting, Interactive Services, News Gathering and other broadband Satellite Applications; Part 2-S2 Extensions (DVB-S2x). DVB Document A171-2
3. JUPITER™ System Bandwidth Efficiency. HUGHES® An EchoStar Company, White Paper, November 2015
4. NEW Jupiter 3/Echostar 24 ultra high density satellite 500 GB. <http://www.dslreports.com/forum/r31558626-NEW-Jupiter-3-Echostar-24-ultra-high-density-satellite-500GB>
5. With Ariane 5 launch of ViaSat-2 and Eutelsat-172b, Arianespace all caught up on protest-delayed missions. <http://spacenews.com/with-ariane-5-launch-of-viasat-2-and-eutelsat-172b-arianespace-all-caught-up-on-protest-delayed-missions/>
6. de Selding, P.B.: ViaSat's-2 'First of its Kind' design will enable broad geographic reach. *Space News* **24**(20), 1–4 (2013)
7. Meet The World's Most Advanced Telecom Satellite. <http://aviationweek.com/connected-aerospace/meet-world-s-most-advanced-telecom-satellite>
8. VIASAT and EUTELSAT to Develop Consumer Broadband by Satellite Services for Europe. http://www.eutelsat.com/news/compress/en/2007/pdf/PR0607_Viasat_Eutelsat.pdf
9. SpaceX Falcon 9 in flawless Inmarsat-5 F4 launch. <https://www.nasaspaceflight.com/2017/05/expendable-falcon-9-inmarsat-5-f4-launch/>
10. Gizinski III, S.J. Manuel, R.: INMARSAT-5 GLOBAL XPRESS®: Secure, Global Mobile, Broadband. https://gobcss.com/wp-content/uploads/2015/06/gizinski.stephen.paper_.pdf.pdf
11. SKY Brasil-1 Satellite. <http://spaceflight101.com/ariane-5-va235/sky-brasil-1/>
12. Capture New Growth with the World's Most Advanced Satellite Platform. <http://www.intelsat.com/media-resources/intelsat-32e-fact-sheet/>
13. Intelsat Epic^{NG} Satellites. The Next-generation, Global High-performance Satellite Platform. www.intelsat.com
14. Operating in an Epic^{NG} Environment. www.intelsat.com
15. Ilchenko, M.Y., Narytnik, T.N., Ye, S., Kalinin, V.I., Cherepenin, V.A.: Environmentally safe communication line with UWB radiation power of 70 nanowatts for wireless local area networks. In: Proceedings of the 21st International Crimean Conference CriMiCo 2011, p. 355–35 (2011)
16. Ilchenko, M.Y., Narytnik, T.N., Kuzmin, S.Y., Fisun, A.I., Belous, O.I., Radzikhovsky, V. N.: Transceiver for 130–134 GHz band and digital radio relay system. *Telecommun. Radio Eng.* **72**(17), 1623–1638 (2013)
17. Ilchenko, M.Y., Narytnik, T.N., Didkovsky, R.M.: Clifford algebra in multiple access noise-signal communication systems. *Telecommun. Radio Eng.* **72**(18), 1651–1663 (2013)
18. Homepage. <http://www.springer.com/Incs>. Accessed 21 Nov 2016



Comparison of Methods for Determining Noise Immunity Indicators of a Multiservice Transmission System

L. A. Uryvsky^(✉) , A. V. Moshynska^(✉) , S. A. Osypchuk^(✉) ,
and B. O. Shmihel^(✉)

National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”, Peremoga Avenue 37,
Kyiv 03056, Ukraine
leonid_uic@ukr.net, avmoshinskaya@gmail.com,
serg.osypchuk@gmail.com, bshmigel@grnail.com

Abstract. The chapter identifies and analyzes the **factors** that affect the **reliability** and, as a result, the **performance** of the communication system. The features of determining the **performance** and **throughput** through **noise immunity indicators** are described. Using several methods for calculating the probability of a symbolic and bit error, data is compared and analyzed. Factors to **improve the noise immunity** of the telecommunication system are justified.

Introduction. Over the past decades, **digital communications** have entered a phase of rapid development, and actively continue developing now. Today, the **wireless broadband networks** are rapidly spreading, potentially transforming from 4G strategies to the latest **5G strategies**. An integral part of these strategies is multiservice, tied to diversity, simultaneous multiplicity, significant volume and high speed of information transmission in the providing of telecommunications services.

The requirements for communications **performance** are the consequence of this. The traditional tool for implementing high performance communication channels is the multi-point modulation (QAM-M: QAM-64, QAM-256, QAM-1024, etc.). But in **wireless broadband communications**, a significant obstacle to achieve the desired performance is the lack of **noise immunity**, which is limited by energy, space dynamics, and other limiting factors.

So, it is **important** to further develop the **fundamental theory** in the field of **information transmission in complex energetically space conditions**, as well as the development of advanced methods at the OSI physical and channel level, such as modulation, coding, etc.

The aim of the work is to research the features of using known and original methods for determining the noise immunity indicators of multi-service information and evaluation of the quality characteristics of information transmission.

The object of the research work is to analyze the methods of estimating the quantitative indices of noise immunity in the high-speed telecommunication system on the physical level of the OSI model.

The subject of the research is the noise immunity of the multi-position signals, combined with a Gaussian noise in the duplex high-speed communication line.

The scientific novelty of the work consists in the development of the vector-phase method (VFM) for accurate analytical determination of the noise immunity indicators for multi-position signals through the reliability indicators of these signals, the feature and advantage of them are their accuracy and complete matching with the simulation modeling results with using the proposed method, as well as quantitative comparisons of formulas that obtained using VFM with known formulas.

Keywords: Throughput · Productivity · Multi-position signals · Signal-code constructions · Frequency · Power · Information efficiency · BPSK · QPSK · PSK · QAM · Vector phase method

1 Introduction

The communication channel capacity is one of the quantitative characteristics that determine a measure of the efficiency of using the resources provided to the telecommunications system at the physical level: frequency and energy.

Multiposition signals are one of the constructive ways to increase the capacity of the communication channels. The essence is the ability to transfer several bits of information in each single manipulated signal, unlike the binary manipulation, where a single signal carries no more than one bit of information.

As it follows from the well-known formula for the bandwidth C of a discrete channel using the manipulation signals of a given multiplicity M [1]:

$$C = V_C \cdot [\log_2 M + (1 - p_s) \cdot \log_2(1 - p_s) + p_s \cdot \log_2(p_s / (M - 1))], \quad (1)$$

On the one hand, the bandwidth increases when the rate of V_C symbols transmission in a channel increases. On the other hand, its growth is limited by the inevitable increase in the probability of error in receiving the p_s symbol while the transmission of V_C symbols accelerates. At the same time, it should be noted that the probability of error p_s is higher, for the higher multiplicity of manipulation M .

Thus, the knowledge of the exact probability value of wrong symbols reception with multiple manipulation p_s is a necessary step in estimating the capacity of the communication channel.

However, the actual amount of transmitted information is determined by the **performance** of the communication channel:

$$R = V_S \cdot [1 + (1 - p_b) \cdot \log_2(1 - p_b) + p_b \cdot \log_2 p_b], \quad (2)$$

where p_b is the probability of the information bit error of the message source, and the transmission rate of the information bits of the message source V_S is related to the transmission rate of the symbols in channel V_C by the relation:

$$V_S = V_C \cdot \log_2 M, \quad (3)$$

because one symbol of a multi-position signal carries $\log_2 M$ bit of the source.

Relations (1)–(3) illustrate, on the one hand, a close relationship between performance indicators and noise immunity, on the other hand, emphasize the relevance of determining the relationship between desired performance R and restrictions on the indicators information transfer p_s and p_b reliability while the source transmission speed V_s and modulation rates M growth.

In the literature [1–7] various formulas are recommended to calculate the probability of error when receiving a symbol of manipulated signals (P_s) and errors when receiving the information bit of the message source (P_{bit}).

Relations obtained by Kantor [2] for signals of multiple Phase-Shift Keying (PSK-M), and formulas for estimating errors when receiving symbols and bits for Quadrature Amplitude Modulation (QAM-M) signals proposed by Prokis [3], should be considered as universal. The universality of formulas is based on a strong argument about the equality of the Euclidean distance between the positions of characters in a multiposition signal. These relations are based on the assumption that the error occurs due to interference as a result of offset of the true position of the character vector on one position. However, this assumption is valid with a significant excess the signal energy over interference, when the probability of vector offset to other positions except neighbor position is extremely small. For small values of the signal -noise ratio, the full group of events should take into account the probability of a signal vector offset to different intervals.

Thus, in addition to the relationships obtained by Prof. Fink [4, 5], the ratios in the mentioned above sources [2, 3] are approximate, i.e. ensure proper accuracy of determining the error probabilities in a certain limited range of values of the energy parameter h_2 , and the probabilities that do not satisfy the physical meaning outside this region (not specified in the sources).

The results obtained by Prof. Fink [4], in turn, refer only to two types of manipulation: BPSK (Binary Phase Shift Keying or PSK-2) and QPSK (Quadrature Phase Shift Keying or PSK-4).

To obtain accurate analytical relationships to determine the probabilities p_s and p_b for a wider range of signal-noise ratios, as well as to justify the methodology for constructing statistical models for calculating the same probabilities with excessive complication of the analytical apparatus, an original method is proposed, based on the physical interaction of the useful signal and interference. In this case, the signal, noise, and their interaction are displayed in vector form, which predetermined the name of the **vector-phase method**.

2 Comparative Analysis of Methods for Determining Noise Immunity for Multiple Phase-Shift Keying Signals

The formula given in the Kantor book [2], is of interest because the analytical dependence is universal for signals of multiple phase manipulations with different positional M , which allows its use for values of $M \geq 2$. Formula

$$P_{b_PSK-M}(h^2) \approx \frac{2}{M} F(\sqrt{h^2} \sin \frac{\pi}{2M}), \quad (4)$$

obtained as a result of approximations of exact formulas, and is the intention to propose a universal formula suitable for calculations.

However, the most accurate and correct method for determining the indicators of noise immunity of multiple phase manipulation signals is the method of determining the probability of error, proposed by L. M. Fink.

Below are the exact formulas proposed by Fink [4], for the bit error probability:

BPSK:

$$P_{b_BPSK}(h^2) = 0,5 \left(1 - F(\sqrt{2h^2}) \right), \text{ where } F(x) = \frac{\sqrt{2}}{\sqrt{\pi}} \int_0^x e^{-\frac{\eta^2}{2}} d\eta \quad (5)$$

QPSK

$$P_{b_QPSK}(h^2) = 0,5 \left(1 - F(\sqrt{h^2}) \right), \text{ where } F(x) = \frac{\sqrt{2}}{\sqrt{\pi}} \int_0^x e^{-\frac{\eta^2}{2}} d\eta \quad (6)$$

PSK-8

$$P_{b_8-PSK}(h^2) = \frac{1}{3} \left(\frac{1}{2} \left(2 - F\left(\sqrt{2h^2} \sin\left(\frac{\pi}{8}\right)\right) - F\left(\sqrt{2h^2} \cos\left(\frac{\pi}{8}\right)\right) \right) \right. \\ \left. + \frac{1}{2} \left(1 - F\left(\sqrt{2h^2} \sin\left(\frac{\pi}{8}\right)\right) F\left(\sqrt{2h^2} \cos\left(\frac{\pi}{8}\right)\right) \right) \right) \quad (7)$$

$$\text{where } F(x) = \sqrt{\frac{2}{\pi}} \int_0^x e^{-\frac{\eta^2}{2}} d\eta. \quad (8)$$

Find the values $h^2 = E_s/N_0$, required to provide $P_b = 10^{-4}/10^{-5}/10^{-6}$, with different methods of calculation.

Figure 1 presents the following results.

1. calculation by the formula of Kantor, given in [2] (lines 1–3);
2. the results given in [2] in Table 4.4 (lines 4–6);
3. calculation by the exact formula Fink [4] (lines 7–9).

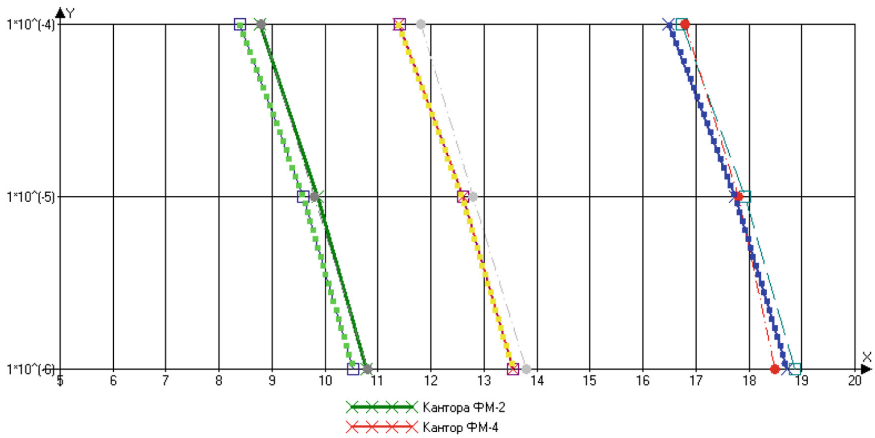


Fig. 1. Graphic representation of the noise immunity of multiple FM signals obtained by various calculation methods.

The tables below present the errors of the values obtained using the Kantor formula (lines 1–3), which differ from the results obtained using the exact Fink formula (lines 7–9) (Table 1):

Table 1. The error values obtained according to the approximate Kantor formula in relation to exact Fink formula

- For BPSK:			
Probability	10^{-4}	10^{-5}	10^{-6}
Difference, h^2	-0,381	-0,279	-0,248
- For QPSK:			
Probability	10^{-4}	10^{-5}	10^{-6}
Difference, h^2	0,0105	-0,0018	0,00131656
- For PSK-8:			
Probability	10^{-4}	10^{-5}	10^{-6}
Difference, h^2	0,0057	0,02845	0,00076

The foregoing leads to the following conclusions.

Formula (4) from [2] is of interest, since the analytical dependence is universal for signals with different positionality M , which allows its use with values $M > 2$. The formula is obtained as a result of approximations of exact formulas, and is an attempt to derive a universal formula suitable for calculations.

So far as the formula is the result of approximations, it has a limited range where its values almost matches with the exact values, i.e.: for QPSK and PSK-8 type manipulations, the range of calculated bit error probabilities is $10^{-4} \dots 10^{-6}$ (lines 2–3).

For the case of BPSK (line 1), the formula (4) in the studied error probability range shows the results with an error of up to 4.5% in comparison with the exact formulas (5)–(7).

At the same time, in other probability ranges, formula (4) is unsuitable, for example:

- for ratio $h^2 = P_s/P_n = 0$, $P_{b_BPSK}(0) = 1$, $P_{b_PSK-8}(0) = 0.33$, while the maximum value of the bit probability is 0.5.

Thus, the universal formula (4) is suitable for use in the range of bit probabilities 10^{-4} – 10^{-6} with an acceptable error for the types of manipulation QPSK and PSK-8.

3 Analysis of the Method of Determining Noise Immunity for Signals of Multiple Quadrature Amplitude

It is known that providing efficient usage of limited energy and frequency resources is a relevant task in telecommunication systems.

Today, QAM is one of the most effective modulation methods, that allows to achieve the highest possible data transmission rates [3, 6, 7].

So, on the one hand, we strive to use the spectrum more efficiently. On the other hand, the higher the multiplicity of manipulation M then the higher the influence from noise on the signal state [3, 7].

The research of the reliability of multiple Quadrature Amplitude Modulation signals, in particular, performed by the method proposed by Prokis [3].

The initial data for the methodology are the following parameters:

h^2 is the energy parameter that characterizes the ratio of the average signal energy E_S at the reception point to the one-sided noise spectral density N_0 :

$$h^2 = \frac{E_S}{N_0} = \frac{P_S}{V_c \cdot N_0}, \quad (9)$$

where V_c – symbol rate in the communication channel, P_S – signal power at the receiving point.

Using these parameters, it is possible to determine the available bit error probability in the channel. It depends on the method of construction and processing the signal.

For QPSK and QAM- M manipulation, the probability of a symbolic and bit error will be determined by the following formulas:

- symbol error probability for QPSK

$$P_{s\ QPSK} = \frac{3}{4} - \frac{1}{2}F(\sqrt{h^2}) - \frac{1}{4}F^2(\sqrt{h^2}), \quad (10)$$

- bit error probability for QPSK

$$P_{b\ QPSK} = \frac{1}{2} \left[1 - F\left(\sqrt{2h^2} \cos \frac{\pi}{4}\right) \right], \quad (11)$$

- symbol error probability for QAM-M

$$p_{s \text{ QAM}} = 1 - \left(1 - \frac{2 \left(1 - \frac{1}{\sqrt{M}} \right)}{\sqrt{2\pi}} \int_{\sqrt{\frac{3}{M-1}h^2}}^{\infty} \exp\left(\frac{-u^2}{2}\right) du \right)^2, \quad (12)$$

- bit error probability for QAM-M

$$p_{b \text{ QAM}} = \frac{4 \left(1 - \frac{1}{\sqrt{M}} \right)}{\sqrt{2\pi}} \int_{\sqrt{\frac{3}{M-1}h^2}}^{\infty} \exp\left(\frac{-u^2}{2}\right) du \quad (13)$$

where M is positioning of modulation.

With the above formulas, were built figures with the dependence of the probability of a **symbolic** and **bit** error on the energy parameters in the communication channel using the **J. Prokis formulas**.

The figures of the symbolic error and the signal-noise ratio for each modulation calculated and built in the Matlab program are as follows:

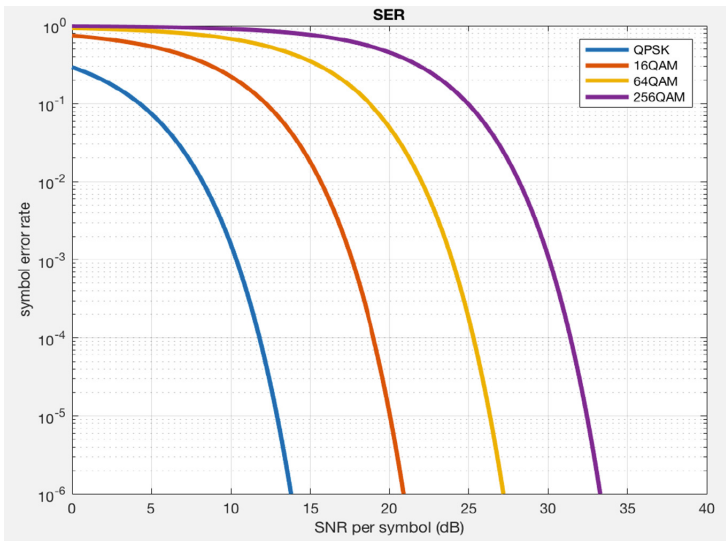


Fig. 2. Symbol errors rates vs h^2 for QPSK, QAM-16, QAM-64 and QAM-256

The bit error probability for an of multi-position signals is less than the symbolic error probability, because not any distortion of a character leads to an incorrect

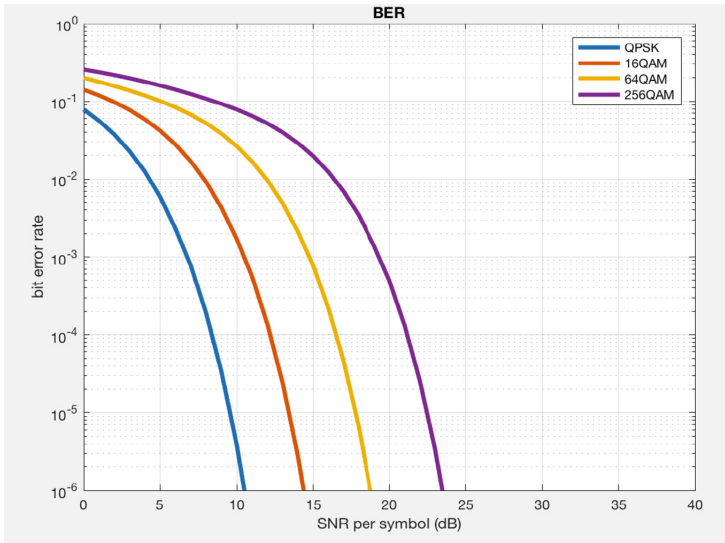


Fig. 3. Bit errors rates vs h^2 for QPSK, QAM-16, QAM-64 and QAM-256

reception of a bit of information. Figure 3 shows the dependence of the bit error probability vs energy parameter h^2 .

From the presented dependencies in Figs. 2 and 3, it can be concluded that in channels with a higher positioning occur a higher number of errors as compared with channels with a lower positioning, hence the Euclidean distance between the characters, which determines the possibility of their difference, is significantly reduced.

4 Analysis of the Vector-Phase Method for Determining Noise Immunity for Signals with Multiple Manipulation

4.1 Problem Statement

The method [8] is based on determining the probability of the result (summary) signal vector and interference in a certain space-phase region, in which the interference leads to incorrect reception of multiple manipulation symbols, and, as a result, to erroneous reception of information bits.

Let the signal S be transmitted to the communication channel. Under the influence of an interference W , the signal S forms the resulting signal r , which arrives to the receiver. If the resulting signal falls to the registration area of the transmitted signal, then the signal is received correctly, if the resulting signal is outside the reliable registration area, then it is not correctly received.

In this case (BPSK signal), the behavior of the interference vector is determined by the relation:

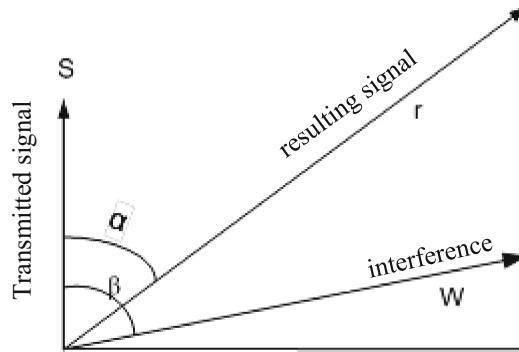


Fig. 4. Vector representation of the interference effect on the resulting signal.

$$W(\beta) = \frac{S}{\sin \alpha(\operatorname{ctg} \beta - \operatorname{ctg} \alpha)} = \frac{S}{-\cos \alpha}, \tag{14}$$

As a result of calculating the probability of interference vector entering the region leading to incorrect reception of an information bit ($\pi/2 < \alpha < \pi$), we get the following formula for BPSK:

$$P_{b \text{ BPSK}}(h^2) = \frac{1}{\pi} \int_{\frac{\pi}{2}}^{\pi} e^{-\frac{h^2}{\cos^2 \alpha}} d\alpha \tag{15}$$

Calculations of the error probability using formula (15) **quantitatively matches** with the results obtained by Fink [4].

4.2 Analytical Model for Determining the Symbol Errors Probability for Multiposition Signals

The next, more efficient type of manipulation than BPSK is QPSK - quadrature phase shift keying. In contrast to BPSK, each transmitted signal with this manipulation transmits 2 bits of information. The view of the QPSK signal space is shown in Fig. 5.

Binary digits in the transmitter are grouped in two, and in each symbol transmission interval two consecutive digits determine which of the four possible signals the modulator will produce. The error probability at the decoder input is determined by the location of the signal points in the ensemble and the noise in the channel. The constellation of signals QPSK is a set consisting of four signal points (S0, S1, S2, S3), offset from each other by 90° and 180°.

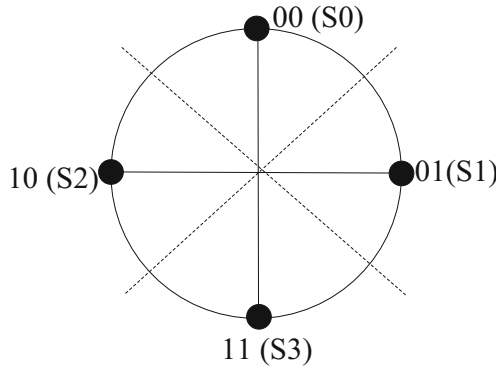


Fig. 5. QPSK signal space using gray code

Consider the case when the symbol S0 corresponds to the correct reception, that is, the vector of signal and interference sum location in the region of the transmitted signal S0. The region of correct signal reception S0 is a beam with an angle of 90°, coming from the origin. Areas of proper reception, for each signal, are shown in dash lines representing the rays. An error will occur during demodulation, if the received signal is in the receiving area of the remaining 3 signals. So, when transmitting S0, an error occurs if the received signal is in the receiving areas S1, S2, S3

Let’s build a vector diagram with relation noise amplitude to its angle, which leads to incorrect reception of the transmitted signal S0 (the resulting sum vector of the transmitted signal and noise will fall into the receiving region of the signals S1, S2, S3). Consider the influence of the noise vector on the signal with an amplitude that equals 1. The vector diagram of the noise amplitudes is shown in Fig. 4. In order to receive the transmitted signal incorrectly, the vector of the resulting signal R, which is formed by the vector of the transmitted signal S and the noise vector W, falls into the reception area of the “erroneous” signal.

Making a series of mathematical transformations over the vectors, we get the following relations between them for the QPSK signals:

$$W(\beta) = \frac{S}{\sin \beta(ctg\alpha - ctg\beta)}, \tag{16}$$

Condition

$$\alpha = \frac{\pi}{4}, \tag{17}$$

for QPSK, describes the boundary state, above which the transmitted signal will be received incorrectly. Using the above condition, we get the boundary condition for the noise vector:

$$W(\beta) = \frac{S}{\sin \beta(ctg\pi/4 - ctg\beta)} \tag{18}$$

The noise vector region described by (18) for the signal S_0 and shown in Fig. 6. The shaded part corresponding to the noise vector that leads to an erroneous reception of the signal S_0 .

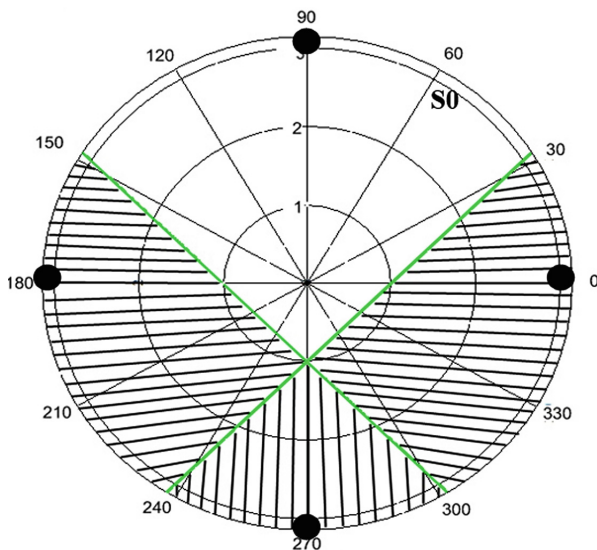


Fig. 6. Vector space of noise signals for QPSK.

Finding the probability of entering the noise vector to the region that will lead to incorrect reception is as follows.

The integral for the wrong reception area “S0” matches with the integral for the wrong reception in areas S1, S2, S3 and looks as follows:

$$\begin{aligned}
 P_{S_QPSK}(S, \sigma) &= \frac{1}{\pi\sigma^2} \int_{\frac{\pi}{4}}^{\pi} \int_{-\frac{S}{\sin\beta}}^{\infty} e^{\frac{-(\eta)^2}{2\sigma^2}} \eta d\eta d\beta = \left| \begin{array}{l} \frac{\eta^2}{2} = t \\ \eta = \pm\sqrt{2t} \\ d\eta = \pm\frac{2}{2\sqrt{2t}} dt \\ \eta = \frac{-S}{\sin\alpha(1-\text{ctg}(\beta))} = \mp\sqrt{2t} \\ t = \frac{S^2}{2\sin^2\beta(1-\text{ctg}(\beta))^2} \end{array} \right| \\
 &= \frac{1}{\pi\sigma^2} \int_{\frac{\pi}{4}}^{\pi} \int_{\frac{S^2}{2\sin^2\beta(1-\text{ctg}\beta)^2}}^{\infty} e^{\frac{-t}{\sigma^2}} \sqrt{2t} \frac{2}{2\sqrt{2t}} dt d\beta = \frac{1}{\pi\sigma^2} \int_{\frac{\pi}{4}}^{\pi} -\sigma^2 e^{\frac{-t}{\sigma^2}} \left| \frac{\infty}{2\sin^2\beta(1-\text{ctg}\beta)^2} \right. d\beta \\
 &= \frac{1}{\pi\sigma^2} \int_{\frac{\pi}{4}}^{\pi} (0 + \sigma^2 e^{\frac{-S^2}{2\sin^2\alpha(1-\text{ctg}\alpha)^2}}) d\alpha = \frac{1}{\pi} \int_{\frac{\pi}{4}}^{\pi} e^{\frac{-S^2}{2\sin^2\alpha(1-\text{ctg}\alpha)^2}} d\alpha
 \end{aligned}
 \tag{19}$$

After replacement $S^2 = P_c$, $\sigma^2 = \frac{N_0}{T}$, $\frac{S^2}{2\sigma^2} = h^2$ we get the symbolic error probability for QPSK:

$$P_{S_QPSK}(h^2) = \frac{1}{\pi} \int_{\frac{\pi}{2}}^{\pi} e^{\frac{-h^2}{\sin^2 \beta(1-\text{ctg}\beta)^2}} d\beta \tag{20}$$

4.3 Analytical Model for Determining the Probabilities of Bit Error Rate (BER) for Multi-position Signals

For QPSK, calculating the bit error rate is reduced to finding the average error probability for each bit. When using Gray codes, the location of the information bits is shown in Fig. 7.

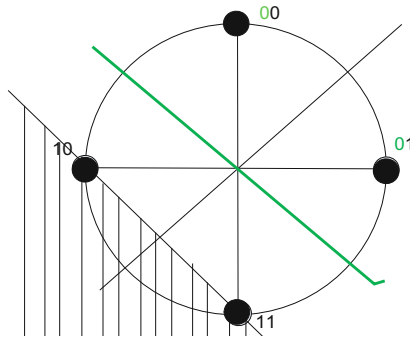


Fig. 7. Constellation diagram for QPSK

If the combinations “10” or “11” are received, an error in the first information bit (when transmitting “0” in the first channel of two) occurs. To place the received signal in the error area, it is necessary that it is affected by the interference, which is described by the following formula:

$$W(\beta) = \frac{S}{\sin \alpha(\text{ctg} \frac{\pi}{4} - \text{ctg} \alpha)} \tag{21}$$

and the noise area, that leads to distortion of the transmitted bit and corresponds to the shaded area in Fig. 7.

Considering the above limitation, calculating the bit error probability for QPSK is reduced to finding the average error probability for each bit, and the result is the following formula:

$$P_{b_QPSK}(h^2) = \frac{1}{2\pi} \int_{\frac{\pi}{4}}^{\frac{3\pi}{4}} e^{-\frac{h^2}{(\sin z - \cos z)^2}} dz \tag{22}$$

For the PSK-8, the bit probability is calculated as the sum of the average error probabilities for each of the three bits that are transmitted in a symbol. The bit errors probability of 1, 2, 3 bits is determined by the probability of a noise vector in the region, which leads to symbol distortion. The average probability of erroneous reception of a single bit for the case of PSK-8:

$$P_{b_PSK8}(h^2) = \frac{2P_{b1_PSK-8}(h^2) + 2P_{b2_PSK-8}(h^2)}{3}, \tag{23}$$

where $P_{b1_PSK-8}(h^2)$ bit error probability of the 1 bit equals to bit error probability of the 3rd bit.

These probabilities correspond to the probability of the noise vector entering the region, that leads to incorrect reception of a bit of information.

So, the error probability of the first bit consists of the probability of hits noise into two areas, shown in Fig. 8.

The probability of noise vector entering this area is calculated as follows:

$$P_{b1_PSK8}(h^2) = \frac{\frac{1}{2\pi} \int_{\frac{\pi}{8}}^{\frac{9\pi}{8}} \int_{\frac{h^2 \tan^2(\frac{\pi}{8})}{(\sin(x) - \cos(x) \tan(\frac{\pi}{8}))^2}}^{\infty} e^{-\eta} d\eta d\alpha + \frac{1}{2\pi} \int_{\frac{3\pi}{8}}^{\frac{11\pi}{8}} \int_{\frac{h^2 \tan^2(\frac{3\pi}{8})}{(\sin(x) - \cos(x) \tan(\frac{3\pi}{8}))^2}}^{\infty} e^{-\eta} d\eta d\alpha}{2} d\alpha, \tag{24}$$

$$P_{b2_PSK8}(h^2) = \frac{1}{2\pi} \left(\int_{\frac{\pi}{8}}^{\pi} \int_{\frac{h^2 \tan^2(\frac{\pi}{8})}{(\sin(x) - \cos(x) \tan(\frac{\pi}{8}))^2}}^{\infty} e^{-\eta} d\eta d\alpha - \int_{\frac{5\pi}{8}}^{\pi} \int_{\frac{h^2 \tan^2(\frac{5\pi}{8})}{(\sin(x) - \cos(x) \tan(\frac{5\pi}{8}))^2}}^{\infty} e^{-\eta} d\eta d\alpha + \right. \\ \left. + \int_{\pi}^{\frac{13\pi}{8}} \int_{\frac{h^2 \tan^2(\frac{5\pi}{8})}{(\sin(x) - \cos(x) \tan(\frac{5\pi}{8}))^2}}^{\infty} e^{-\eta} d\eta d\alpha - \int_{\pi}^{\frac{9\pi}{8}} \int_{\frac{h^2 \tan^2(\frac{\pi}{8})}{(\sin(x) - \cos(x) \tan(\frac{\pi}{8}))^2}}^{\infty} e^{-\eta} d\eta d\alpha \right) \tag{25}$$

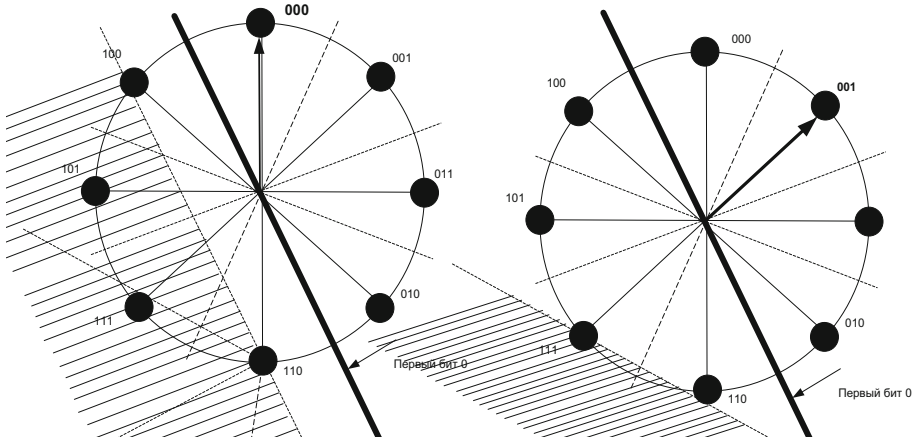


Fig. 8. Noise regions for the first bit of the PSK-8 signal

The error probability of the third bit is determined by the noise vector area and it is shown in Fig. 9.

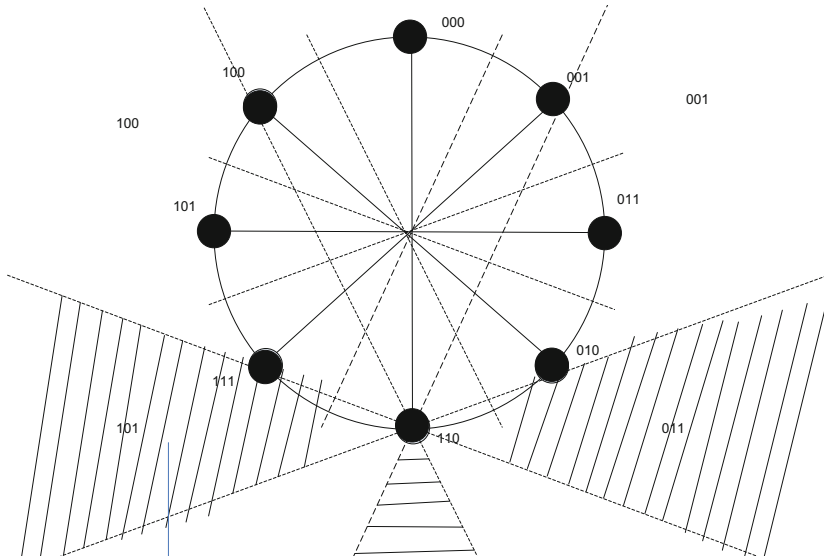


Fig. 9. Noise regions for the third bit of the PSK-8 signal

For the case of QAM-16 (Quadrature Amplitude Modulation - 16), the probability of a bit error is also determined as the average probability of a bit error across all 4 bits. In this case, the average bit probability of the first bit error is equal to the error probability of the third bit, and the bit error probability of the second bit matches with

the probability of the fourth bit error. The equality of bit probabilities is provided by the similarity of the corresponding reception areas.

Since the probabilities of errors 1 and 3, 2 and 4 bits match, therefore, to find the average bit error, it is enough to find the probabilities of errors 1 and 2 bits:

$$P_{b_QAM16}(h^2) = \frac{P_{b1_QAM16}(h^2) + P_{b2_QAM16}(h^2)}{2} \tag{26}$$

where $P_{b1_QAM16}(h^2)$ is bit error probability of the 1 and 3 bit, $P_{b2_QAM16}(h^2)$ is bit error probability of the 2 and 4 bit.

The error probability of the 1st bit is determined by the conditions for the noise vector to enter the region shaded in Fig. 10.

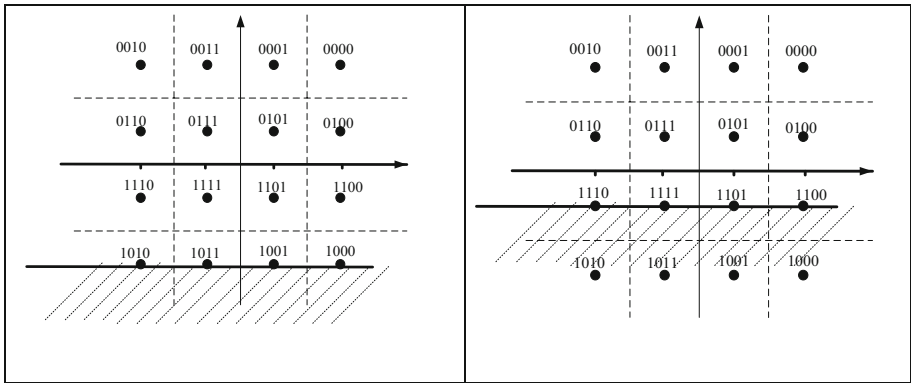


Fig. 10. Areas of interference for the 3 bit for the QAM-16 signal: (a) for signals 0010,0011,0001,0000, (b) for signals 0110,0111,0101,0100

These probabilities are calculated as follows:

$$P_{b1_QAM16}(h^2) = \frac{\frac{1}{2\pi} \int_{\frac{\pi}{4}}^{\frac{5\pi}{4}} \frac{-h^2}{e^{2(\cos z - \sin z)^2}} d\alpha + \frac{1}{2\pi} \int_{\arctg(3)}^{\pi + \arctg(3)} \frac{-5h^2}{e^{18(\cos z + \frac{\sin z}{3})^2}} d\alpha + \frac{1}{2\pi} \int_{\frac{\pi}{4}}^{\frac{5\pi}{4}} \frac{-h^2}{e^{18(\sin z - \cos z)^2}} d\alpha}{4}$$

$$P_{b2_QAM16}(h^2) = \frac{\frac{1}{2\pi} \int_{\frac{\pi}{4}}^{\frac{5\pi}{4}} \frac{-h^2}{e^{18(\cos z - \sin z)^2}} d\alpha - \frac{1}{2\pi} \int_{\frac{\pi}{4}}^{\frac{5\pi}{4}} \frac{-h^2}{e^{50(\cos z - \sin z)^2}} d\alpha + \frac{1}{2\pi} \int_{-\frac{3\pi}{4}}^{\frac{5\pi}{4}} \frac{-h^2}{e^{18(\cos z - \sin z)^2}} d\alpha}{2} \tag{27}$$

A feature of the formulas that were obtained using the vector-phase method is their accuracy. It does not contain any assumptions about universalization at the expense of the calculations reliability.

The calculations by formulas (14), (16), (18), (21) numerically match with the known results [4]. Relations (20), (21), (23) ... (27) are the product of an exclusively vector-phase method for determining the error probability.

The results obtained using the proposed method are correct over the entire range of parameters studied, while widely used approximate formulas are sufficiently accurate only in a narrow range of the initial values (for high-order manipulation and with significant energy in the communication channel).

At the same time, the accuracy of analytical calculations within the boundaries of the VFM is provided by the complex calculations of numerical integration, although it does not contain assumptions about their approximation.

5 Comparative Analysis of Methods for Determining Noise Immunity for Signals of Multiple Manipulation

To compare the vector-phase method with some other known formulas, it is necessary to analyze and compare to get data.

The results comparison from obtained formulas with known ones (Fig. 11) shows complete matching of the results for the case of BPSK and QPSK with previously obtained by outstanding scientists that were working in the noise immunity of multiposition signals field [4, 9]. It confirms the accuracy and adequacy of the proposed method for determining the error probability for the multiposition signals. At the same time, for signals, for which are not given analytical expressions in the well-known works of scientists [2, 6], the differences in the results were up to 0.2 dB on the h^2 scale for PSK-8, and for QAM-16 signals the difference from the known approximate expressions is up to 4 dB.

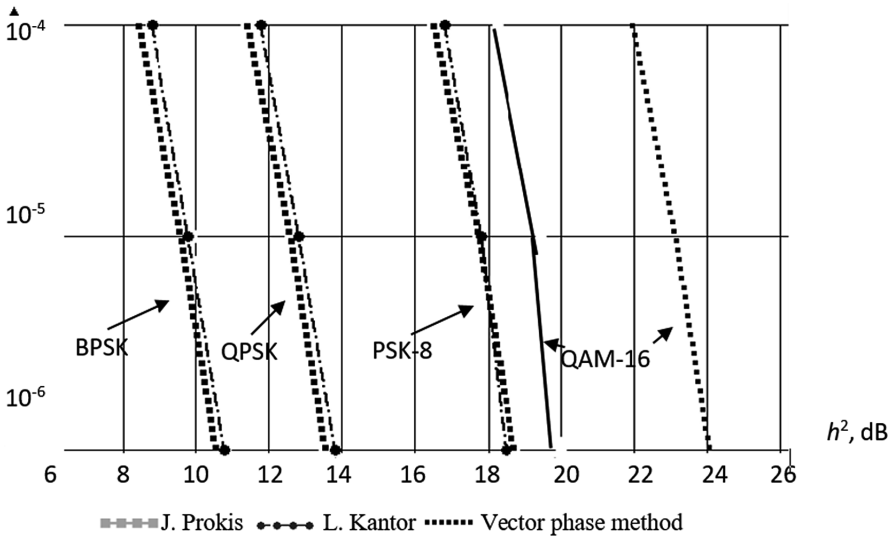


Fig. 11. Comparative analysis of the known methods for estimating the noise immunity of multi-position signals.

Consequently, due to the vector-phase method, the exact analytical relations constructed to determine the bit and symbols error probability of multi-positional manipulation. They are correct for any channel energy.

A significant disadvantage of the vector-phase method is its complexity at the stage of numerical integration. Therefore, an alternative to analytical calculations is the usage of simulation techniques within the borders of the proposed method. In particular, for the QAM-64 signals, the calculation of reliability indicators is performed in this way.

So, the usage of the vector-phase method for calculating the signals error probability for multiple manipulations based on simulation modeling allows to simplify the calculation procedure by eliminating complex analytical relationships for high orders of manipulation (for example, QAM-64).

An example of the relevant usage of the obtained quantitative noise immunity indicators is the task of determining the compliance of the signal-to-noise power ratio intervals and the recommended types of multi-position manipulation by the criteria of achieving the highest communication channel capacity according to the expression (1).

The relation of the communication channel bandwidth and its energy state is shown in Fig. 12.

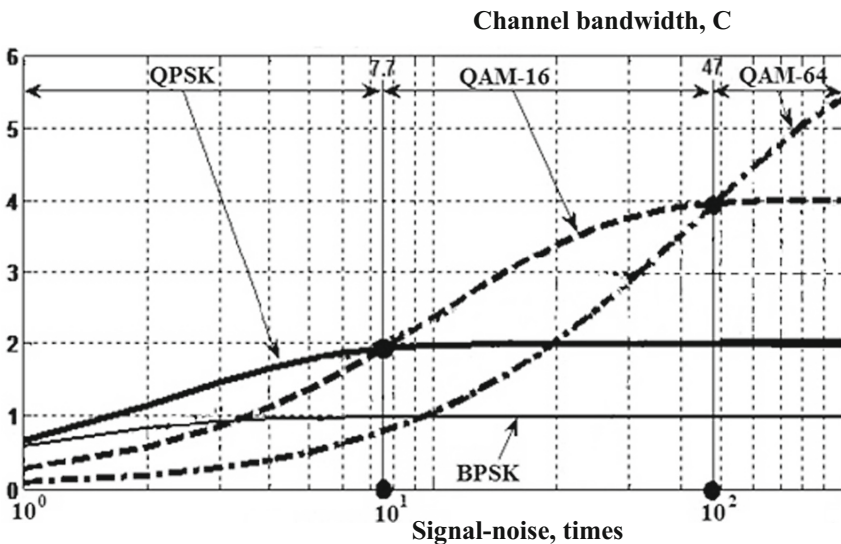


Fig. 12. The relation of the communication channel bandwidth to signal-noise ratio for different multi-modulation types

It can be concluded that, according to Fig. 12, it is advisable to use modulation types considered by the criterion of maximum throughput depending on the signal energy to noise spectral density ratio at the receiving point (Table 2).

Table 2. Signal/Noise limit values for selected types of multi-position modulation

Interval №	Signal-noise ratio, times	The optimal type of multi-position modulation by bandwidth criterion
1.	$0 < h^2 < 7,7$	QPSK
2.	$7,7 < h^2 < 47$	QAM-16
3.	$47 < h^2 < 100$	QAM-64

Since manipulations of higher orders worsen the noise immunity of telecommunication channels, for wireless information transmission systems it is proposed to use a well-known way to deal with this phenomenon—adaptive manipulation.

6 Conclusion

1. Modern society seeks to satisfy its own needs from the side of receiving information with high speed transmission and reliability, therefore the problem is quite acute in the conditions of limited communication channels resources, namely frequency and energy [9]. Scientists and engineers are faced with the task of the most efficient usage of these resources and keep the indicators of the main channel parameters [10].

Multiposition signals are one of the constructive ways to increase the communication channels capacity. The essence is the ability to transfer several bits of information in each single manipulated signal.

Relations (1)–(3) for the communication channel capacity and performance, emphasize the relevance of determining the relationship between desired performance R and restrictions on the indicators information transfer p_S and p_b reliability while the source transmission speed V_S and modulation rates M growth.

The article analyzes the influence of positional manipulation to the noise immunity of discrete telecommunication channels. It is shown that manipulations with higher orders worsen the noise immunity of telecommunication channels. For wireless information transmission systems, it is proposed to use a known way to deal with this phenomenon - adaptive manipulation.

2. Relations obtained by Kantor [2] for multiple Phase-Shift Keying (PSK-M) signals, and formulas for estimating symbols and bits errors for Quadrature Amplitude Modulation (QAM-M) signals proposed by Prokis [3], should be considered as universal.
3. The results comparison from obtained formulas with known ones (Fig. 11) shows complete matching of the results for the case of BPSK and QPSK with previously obtained by outstanding scientists that were working in the noise immunity of multiposition signals field [4]. It confirms the accuracy and adequacy of the proposed method for determining the error probability for the multiposition signals. At the same time, for signals, for which are not given analytical expressions in the well-known works of scientists [2, 6], the differences in the results were up to 0.2 dB on the h^2 scale for PSK-8, and for QAM-16 signals the difference from the known approximate expressions is up to 4 dB.

4. The developed vector-phase method is new for accurate determination of noise immunity of multi-position signals. A feature and advantage of the formulas obtained using the vector-phase method is their accuracy and complete matching with the results of simulation using the proposed method.

The results obtained using the proposed method are correct over the entire range of parameters studied. In the same time, widely used approximate formulas are sufficiently accurate only in a narrow range of the initial values (for high-order manipulation $M \geq 16$ and with significant energy in the communication channel $h^2 > 10$).

The accuracy of analytical calculations within the boundaries of the VFM is provided by the complex calculations of numerical integration, although it does not contain assumptions about their approximation.



The usage of the vector-phase method for calculating the signals error probability for multiple manipulations based on simulation modeling. It allows to simplify the calculation procedure by eliminating complex analytical relationships for high orders of manipulation (for example, QAM-64).

References

1. Gallager, R.G.: Information Theory and Reliable Communication/Massachusetts Institute of Technology. Wiley, New York (1968). 720 p
2. Kantor L.Y.: Satellite communication and broadcasting. Radio and Svyaz, Moscow (1988)
3. Prokis, J.G.: Digital Communication, 3rd edn. McGraw-Hill, New York (1995)
4. Fink, L.M.: The Theory of the Transmission of Discrete Messages. 2nd edn. Publishing House Soviet Radio (1970). 728p
5. Korzhik, V.I., Fink, L.M., Schelkunov, K.N.: Calculation of noise immunity of systems of discrete message transmission. Radio and Svyaz, Moscow, 231 p. (1981)
6. Chiani, M., Win, M.Z., Zanella, A.: Error probability for optimum combining of M-ary PSK signals in the presence of interference and noise. IEEE Trans. Commun. **51**, 1949–1957 (2003)
7. Banket, V.L.: Signal-code constructions in telecommunication systems. Banket V.L., Odessa, 180 p. (2009)
8. Uryvsky, L.A., Prokopenko, E.A., Natalenko, A.I.: Vector-phase method for determining the noise immunity indicators for multi-position signals. *Zv'yazok* **4**, 58–68 (2012)
9. Ilchenko, M.E., Moshinskaya, A.V., Uryvsky, L.A.: Levels separation and merging in the OSI reference model for information-telecommunication systems. *Cybern. Syst. Anal.* **47**(4), 598–605 (2011). <https://doi.org/10.1007/s10559-011-9340-4>
10. Ilchenko, M.: Aspects of system analysis in the applied information theory for telecommunications. In: Ilchenko, M., Uryvsky, L. (eds.) *Cybernetics and Systems Analysis*, vol. 46, no. 5, pp. 737–743 (2010). <https://doi.org/10.1007/s10559-010-9255-5>



Research of the Control Algorithms for a State of Duplex Communication Channel in the Conditions of Multipath

L. Uryvsky^(✉) , S. Osypchuk^(✉) , and V. Solyanikova^(✉)

National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”, Peremoga Avenue 37,
Kyiv 03056, Ukraine
leonid_uic@ukr.net,
serg.osypchuk@gmail.com, valeriasoljanikova@gmail.com

Abstract. This chapter is devoted to the features of the modern DVB-T2 standard technology research, which are designed to provide high noise immunity to the information transmission in duplex multiservice high data rate channels. The object of the study is the influence of transformations used in the DVB-T2 systems on the communication line noise immunity. The expediency of borrowing and transferring the tools that used in DVB-T2 into the duplex communication line is considered, in order to increase the noise immunity of the system. The proposals on increasing the noise immunity of the duplex communication line with the use of DVB-T2 standard transformations are given. In this paper, the parameters calculation of a typical radio relay link is carried out, and the adaptive encoding and modulation scheme for a duplex transmission system is developed based on a DVB-T2 MCS set. The channel models with fading are analyzed and the choice of the channel model and methods of calculating the noise immunity parameters using the DVB-T2 technology are substantiated.

Keywords: Duplex communication line · DVB-T2 standard · Noise immunity · Modulation · Encoding · Modulation-coding scheme · Radio relay line · Fading

1 Introduction

Digital communications are currently actively developing worldwide – this is the main trend in telecommunications [1]. The major directions in the transmission systems development are the communication lines efficiency improving, communication range increasing, quality and reliability improving, continuous improving of technical elements and equipment. Digital transmission systems include all of these components.

The DVB project consortium has begun developing standards for digital television since 1991. Standards were names as DVB (Digital Video Broadcasting).

The most widely adopted and expanded digital terrestrial television standard is DVB-T, which was published in March 1997. This system is most common in the world and contains modern modulation and coding types, which enables highly

efficient use of valuable terrestrial range for audio, video and data services for fixed, portable and mobile devices.

DVB consortium released a new standard for terrestrial digital broadcasting in 2008: DVB-T2 [2–4]. The DVB-T2 standard is the improved and functionally advanced successor standard of DVB-T. It retained the main ideas of the signal processing (scrambling, data interleaving, encoding), but each stage is enhanced and expanded. DVB-T2 modification is the perfect solution that provides high signal stability and the necessary capacity increase.

The peculiarity of digital terrestrial television signal transmission is the **uncertainty** of the conditions that the reception point will be in. The channel is not only **vulnerable** but it is also in a high building concentration area where there are much **industrial noises** with a high risk to cause the **multipath propagation**. Therefore a noise immunity should be high enough to ensure the high signal reception **reliability**.

Thus, the **purpose** of this study is to examine the **DVB-T2** standard features, designed for **high noise immunity** information transfer and these features usage in a **high data rate duplex multiservice communication channel**.

The object of study is the impact of transformations used in TV system DVB-T2 to communication line noises immunity.

The subject of study is the research all of these transformations in the duplex high data rate lines that are susceptible to multipath propagation.

According to the work purpose, the main **tasks** are:

1. To explore the technological **features** of DVB-T2 standard that assure **noise immunity** in a **duplex** transmission system that is susceptible to multipath propagation.
2. To prove a **feasibility** of using **DVB-T2** technologies and research their application in a high data rate duplex communication system (**HDRDCS**).
3. To analyze the channel models with **fading**, justify the channel model **choice** and **noise immunity** calculation **method**.
4. To **suggest** the **noise immunity** improvements for duplex links using the known technologies from DVB-T2.

2 Study of DVB-T2 Technological Features that Assure Improving the Noise Immunity in HDRDCS

A HDRDCS implements a duplex communication and may both transmit and receive information at any time. It uses both noise immune coding and physical signal shape for its formation. A classic example of duplex communication is the radio relay communication [5].

The notable feature of digital terrestrial television signal transfer is an uncertainty of reception point conditions. The channel is not only vulnerable but it is also in the residential area where there are much industrial noises that causes a high risk of multipath propagation. Therefore a noise immunity should be high enough to ensure high signal reception reliability.

Resources of terrestrial digital television are limited while the users number is growing continuously. Therefore, the need exists to increase the communication channel resources usage efficiency with keeping constant such characteristics of communication as data rate and distance.

In general, the throughput of DVB-T2 system depends on the following parameters:

- Modulation type (QPSK, QAM-16, QAM-64, QAM-256),
- LDPC and BCH codes parameters [6],
- OFDM parameters [7],
- FFT dimension (Fast Fourier Transformation size),
- Guard interval length (1/4, 19/256, 1/8, 19/128, 1/16, 1/32, 1/128), etc.

The DVB-T2 system can transmit multiple independent media streams, each with its own modulation scheme, encoding rate, and time intervals.

The task complexity of transferring many services to many users requires managing multiple service and defense signal transformations associated with the signals addition and redundant symbols. Hence, the actual symbol rate of DVB-T2 is significantly higher than the real rate of information flow. This again triggers a system to have the redundant noise immunity. This is the price for multiservice provisioning in the DVB-T2 standard [2–4].

A use of the **BCH/LDPC** correcting codes in DVB-T2 provides a much higher effectiveness than traditional error correction codes provide, what gives a much higher coding rate and increases significantly the overall channel bandwidth.

It is known that DVB-T2 standard is the example of modern multiservice multi-thread system features implementation with reliable information transmission to many users, but in the unidirectional communication mode. **Promising** to be considered is the task of implementing multiservice duplex communication systems based on the set out principles in DVB-T2 standard, specifically for radio relay communication systems.

Thus, DVB-T2 components features that are designed for high noise immunity information transmission, can be considered for high data rate duplex communication systems. Taking an advantage of duplex system, it provides the opportunity to make the right choice about combinations of parameters described such as modulation type and noise immune code, in line with minimizing losses in a given coverage area, optimizing the transmitter power value, bandwidth and system stability.

2.1 Providing High Noise Immunity in HDRDCS

Noise immunity is the ability of communication system to counteract the harmful effects of noises.

Different sorts of noises affect the transmitted signals over the air. The origin of noises can be divided into two classes: artificial and natural. The internal heat radio noises, thermal radiation noises of the Earth and atmospheric noises attribute to the natural origin noise types [1].

The thermal noise energy is uniformly distributed in a wide frequency range, so the noise spectrum is uniform. The thermal noise because of these properties is called “white noise”. White noise power at the output of bandwidth Δf determined by the expression (1):

$$P_W = N_0 \cdot \Delta f, \quad (1)$$

where N_0 – noise power density per 1 Hz of frequency band.

The additive and multiplicative interferences are differentiated depending on the noise effect on the signal. The additive noises are caused by presence of third-party radio emissions of various origins: natural (the natural emitters) and artificial (other radio channels). The multiplicative noises are often called “fading” and they are implied by radio transmission states (named as slow fading) and multipath radio propagation phenomena (named as fast fading) [1, 8].

The useful signal power should be much higher than noise power to ensure a stable reception: $P_s \gg P_n$ [1].

An easy way to improve noise immunity is to increase the SNR value by increasing the transmitter power, but this method may not be appropriate because of the existing constraints on radiated power and due to a significant complexity increase and equipment cost.

The important way to improve a signal noise immunity is a rational choice of modulation type. A significant increase of signal noise immunity and significant signal bandwidth increase can be achieved when applying different modulation types.

The excessive way to improve noise immunity is the use of noise immunity codes.

The noise immune reception approach consists in using a redundancy and a priori information about signal interference to solve problems of receiving signal in the optimal way: identifying signal, differentiating signals or messages recovery by errors correction.

2.2 Analysis of DVB-T2 Technological Features for Improving Noise Immunity in HDRDCS

The DVB-T2 standard is improved and is functionally advanced successor of DVB-T standard. It retained the main ideas of signal processing (scrambling, data interleaving, encoding), but each stage is enhanced and expanded. Figure 1 shows the general DVB-T2 signals processing scheme.

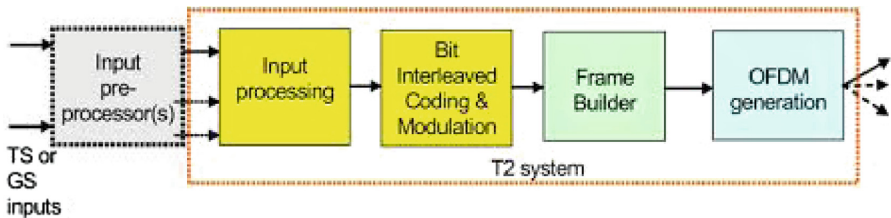


Fig. 1. The DVB-T2 general scheme of processing transmitted signals

The redundant noise immunity encoding mechanism is the one of DVB-T2 main features. The antinoise robust code is the concatenated one: the external code is the Bose-Chaudhuri-Hocquenghem (BCH), and the inner code – the code with a low

density parity checks (LDPC). Before modulation (unless BPSK and QPSK cases), the code words are subjected bitwise by interleaving and distributing to modulation symbols.

The modulation QAM-256 is added in DVB-T2. It increases the transmission channel capacity by 33% (compared to QAM-64 modulation in DVB-T [2–4]). Typically, the transition from QAM-64 to QAM-256 requires SNR increasing by 4–5 dB. However, DVB-T2 encoding rate may be much higher and the overall bandwidth significantly increases due to the use of BCH and LDPC corrective codes whose effectiveness is much higher than traditional error correction codes (including Reed-Solomon).

The DVB-T2 standard has also changes in the OFDM symbols structure. The nominal possible subcarriers number increased and the other modes added: 1K, 4K, 16K and 32K subcarriers.

2.3 Peculiarities of DVB-T2 Signal Processing Features for Improving a Noise Immunity of HDRDCS

The main technologies to improve a noise immunity in the DVB-T2 standard are the following:

- Concatenated coding;
- High index modulation types;
- Interleaving (bits, cells, time, frequency);
- OFDM [7].

This provides the following data rates for the selected modulation-coding schemes (MCS) (Table 1).

Table 1. Maximum bits stream rates at 8 MHz bandwidth, 32 K nominal subcarriers, and guard interval 1/128, scheme pilot subcarriers RR7

Modulation	Coding rate	The absolute maximum rate, mbps	Modulation	Coding rate	The absolute maximum rate, mbps
QPSK	1/2	7.49255	QAM-64	1/2	22.51994
	3/5	9.003747		3/5	27.06206
	2/3	10.01867		2/3	30.11257
	3/4	11.27057		3/4	33.87524
	4/5	12.02614		4/5	36.1463
	5/6	12.53733		5/6	37.68277
QAM-16	1/2	15.03743	QAM-256	1/2	30.08728
	3/5	18.07038		3/5	36.15568
	2/3	20.10732		2/3	40.23124
	3/4	22.6198		3/4	45.25828
	4/5	24.13628		4/5	48.29248
	5/6	25.16224		5/6	50.34524

The BCH and LDPC codes provide the best coding correction what allows to transmit more data over a dedicated channel; they also have a better BER slope as a function of SNR [1, 6].

Another notable feature of DVB-T2 standard is the variety of interleaving types. There are four different interleaving types [2–4]:

- Bits interleaving;
- Cells interleaving;
- Time interleaving;
- Frequency interleaving.

The OFDM parameters selection, namely the guard intervals, pilot signals and FFT dimension, MCS types, causes an increased bandwidth and better system immunity [7].

2.4 Proposals for Improving DVB-T2 Signal Processing Procedures in HDRDCS

High modulation types and coding sets (up to a QAM-4096) are used in duplex transmission systems, including radio relay lines. The use of such a big number of MCS leads to more frequent switching between MCS modes. In addition, the smaller modes used, the less CPU overhead and simpler scheme is. It is known that a higher the modulation type is, the bigger potential errors probability is expected due to an interference or absorption.

Therefore, it is appropriate to consider proposals concerning the MCS sets reduction in a duplex mode line relative to the possible MCS combinations used in DVB-T2, and choose the most optimal MCS combinations set for HDRDCS.

3 Fading Channel Models Analysis for HDRDCS

3.1 Factors that Affect the Waves Propagation in Free Space

Radio waves propagation usually depends on several factors, regardless of what radio communication serves for or what is the radio communications goal.

Radio waves propagation waves depends on the frequency band. The terrain and meteorological troposphere characteristics are prevailing impact at microwave frequencies mainly. However, the gas and precipitation absorption effects should take into account at frequencies above 6 GHz. The rain effects predominant at frequencies close to 10 GHz. The gas absorption begins to affect at the frequency about 22 GHz, where the water vapor shows a peak characteristic.

The “multipath propagation effects” term is applied to those cases where the effective received signal consists of several components that arrive at the receiving antenna in different ways.

The multipath propagation components can have different phases and different amplitudes, and their mutual relation may also continuously change over time. The multipath propagation effects appear due to microwave reflection from buildings, from the Earth’s surface or in a result of reflection from horizontal boundaries between the

different atmosphere layers. The multipath propagation effects that are created due to waves reflections, cause the fast fading that is observed in radio waves. Such fast fading may seriously degrade the QoS.

Radio waves can propagate differently, typically causing attenuation, depending on a topography and weather conditions. One of the main radio assessment objectives is fading between transmitters and receivers. The propagation mechanisms classification is the following on this matter:

- Propagation in free space;
- Refraction;
- Diffraction;
- Reflection and scattering;
- Absorption;
- Fading in the rain.

The **fading** may occur due to an **interference** between the component fields of sight and components reflected from the earth (e.g. from atmospheric layers and buildings), when the line of sight is significantly far above the Earth's surface; while **diffraction** losses may be **avoid** at the same time due to line of sight is significantly far above the Earth's surface.

The multipath propagation effects can lead to short fading. In addition, fading due to absorption in atmospheric gases and rain can be even more important at frequencies above 10 GHz. The rain attenuation effects can lead to signal fading with more significant time length.

The relative importance of fading due to rain and multipath propagation effects depends on the frequency, climate and path length.

It was found that the best channel is formed in the countryside areas where there are no sources for re-radiation. The city buildings model causes the urban multipath propagation due to reflections from buildings. The mountains and hilly areas conditions are the worst for digital communication signal propagation.

3.2 The Channel Fading Model Selection for HDRDCS

Currently, there are many options for the mathematical description of wireless communication channels, such as Lee model, Khata model, Okamura model, and others. But calculations should be performed in a model that describes a channel with fading. For this case, the suitable are Rayleigh [9] and Rice [10] models that describe fading in multipath propagation channels. Let's consider both models and choose the one that meets these criteria.

The Rayleigh and Rice fading channels are widely used for modeling the physical phenomena in wireless communication systems, including the effects of multipath scattering, time dispersion and Doppler shifts resulting from the relative motion between transmitter and receiver [8].

Figure 2 shows the direct beam and the main path of the reflected rays between stationary transmitter and receiver. The shaded figures are reflectors such as buildings.

The major signal part comes to the receiver with a delay. In addition, the radio signal undergoes dispersion in a local scale for each major way. This local dispersion is

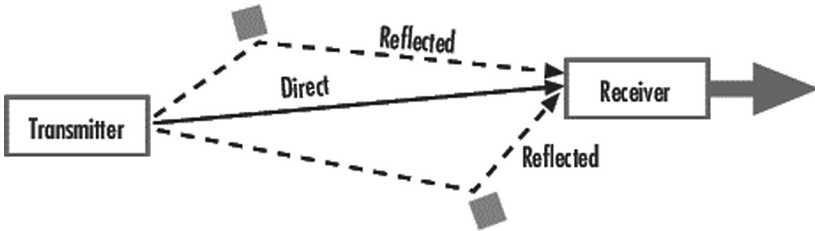


Fig. 2. Rays path between stationary radio transmitter and receiver

usually characterized by multiple reflections from objects near the receiver. These components are combined in the receiver and cause a phenomenon known as multipath fading. Because of this phenomenon, each propagation way is a discrete fading path. Typically, such fading process is characterized by the Rayleigh distribution for the missing line of sight and by the Rice distribution for the line of sight [8].

The Rayleigh channel describes a channel with fading when its level varies randomly, and the carrier signal phase is distributed by uniform distribution law.

The signal amplitude is described by Rayleigh distribution in such channel. Rayleigh distribution – this is the random variable X probability distribution with a density as per the formula [9]:

$$f(x, \sigma) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right), x \geq 0, \sigma > 0, \tag{2}$$

where σ is the scale parameter. The corresponding distribution function is:

$$P(X \leq x) = \int_0^x f(\varepsilon) d\varepsilon = 1 - \exp\left(-\frac{x^2}{2\sigma^2}\right), x \geq 0. \tag{3}$$

Figure 3 shows the probability of Rayleigh density distribution function.

The Rice channel is a channel with fading where the signal amplitude distribution is a decomposition of two independent components: the constant one that describes the direct beam channel without fading, and the variable one that characterizes the fading signal (this component usually represents the multiple reflected rays) [10].

An amplitude of the total signal is random and obeys the Rice distribution. The Rice distribution is a generalization of the Rayleigh distribution.

If X and Y are the independent random variables with the normal distribution with equal σ^2 variances and non-zero mathematical expectation (in general, not equal), the value $Z = \sqrt{X^2 + Y^2}$ subordinated the Rice distribution, and the probability density is determined as [10]:

$$f(x|v, \sigma) = \frac{x}{\sigma^2} \exp\left(-\frac{(x^2 + v^2)}{2\sigma^2}\right) I_0\left(\frac{xv}{\sigma^2}\right), \tag{4}$$

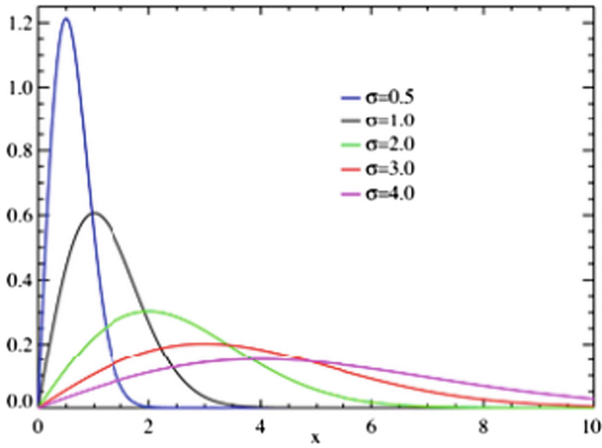


Fig. 3. The distribution function of Rayleigh probability density

where $I_0(z)$ is the modified Bessel function of the first kind zero order, $v = \sqrt{\mu_1^2 + \mu_2^2}$, μ_1 and μ_2 are the mathematical expectations of X and Y .

Figure 4 represents the probability density function of Rice distribution.

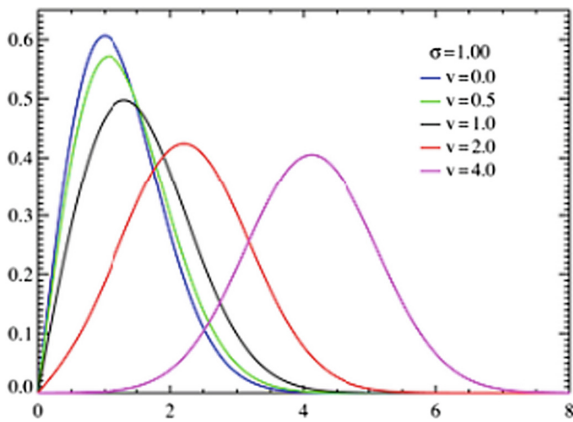


Fig. 4. The distribution function of Rice probability density

Let's compare the signal distribution channels for both models (Table 2).

Table 2. The signal paths comparison for Rayleigh and Rice channels

Signal distribution	Channel with fading
The line of sight from the transmitter to the receiver	Rice channel
One or more major reflected rays from the transmitter to the receiver	Rayleigh channel

It can be concluded that the closest to the radio relay link is the Rice channel model. The Rice channel has a main beam and many reflected beams with relatively weak energy, while the Rayleigh channel doesn't have a main beam and has multipath beams with equal energy.

4 Proposals to Improve a Noise Immunity in HDRDCS by Using the Techniques from DVB-T2 Standard

4.1 Data Rate Dependency Analysis as a SNR Function with Using the Adaptive MCS, and Proposals for MCS Set Reduction

Let's consider options for adaptive coding and modulation (MCS) for radio relay link. Since one of the main research tasks is to analyze the possibilities of using DVB-T2 features for HDRDCS, let's take the DVB-T2 MCS set (Table 1) as the input modes with respective modulation and coding parameters.

Table 1 gives a bit rate that can be achieved with the requested parameters when using MCS modes from DVB-T2 standard. Let's construct the BER dependency of signal energy that falls on a received 1 bit E_b , to the noise power spectral density N_0 for each MCS mode (Fig. 5).

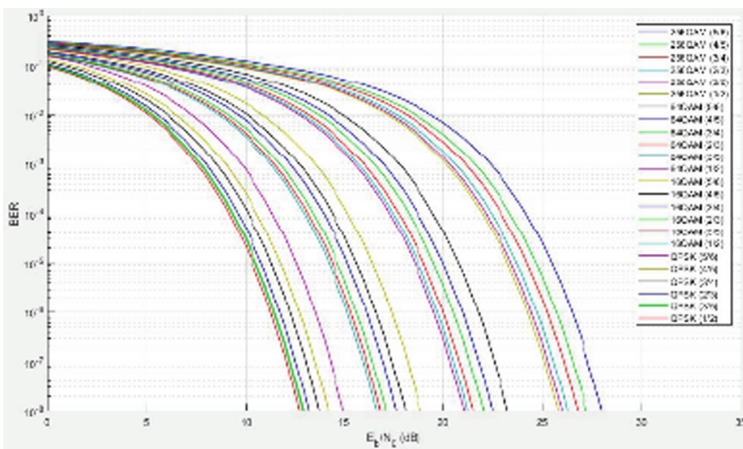


Fig. 5. BER dependency on E_b/N_0 for DVB-T2 standard MCS modes

As seen from Table 1 and from Fig. 5, several MCS modes have equal or very close bitrate values, and therefore it is advisable to use only those MCS modes that are more noise immune, or in other words – can operate at lower E_b/N_0 values. Additionally, it makes sense to do not use the MCS modes that are visually close to each other, because using adaptive MCS switching between them are performed frequently and not always appropriate due to increased system complexity and functionality.

In view of the above, let’s reduce the MCS modes number for a HDRDCS (Table 3). The reduction will optimize an adaptive MCS switching process, reduce the CPU load and simplify the MCS selection decision-making algorithms during the transition to a particular MCS mode.

Table 3. The result of reduction the DVB-T2 MCS modes set

	5/6	4/5	3/4	2/3	3/5	1/2
QAM-256	+	-	+	-	-	-
QAM-64	+	-	+	-	-	-
QAM-16	+	-	-	+	-	-
QPSK	+	-	-	-	-	+

The MCS reduction result is graphically shown on Fig. 6.

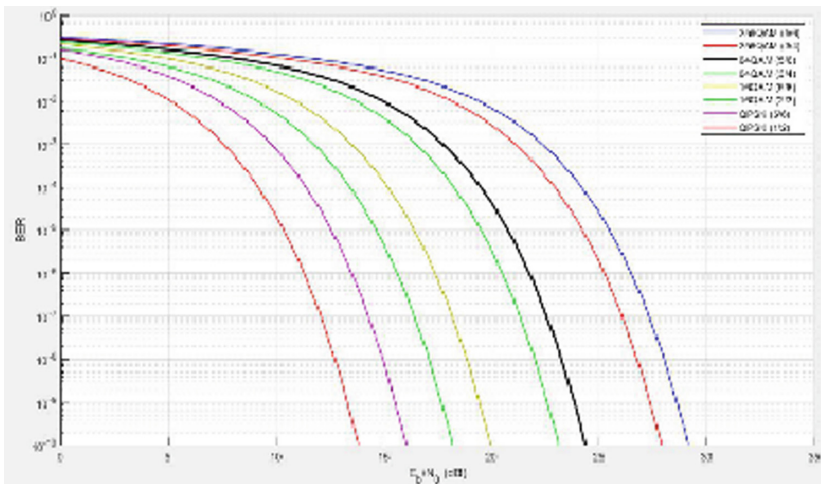


Fig. 6. The graphical representation of BER dependency on E_b/N_0 for MCS set after DVB-T2 MCS modes number reduction

Let’s find the thresholds of E_b/N_0 (h^2) at what the MCS modes switching will be performed to MCS modes with higher noise immunity and lower bit rate if fading occurs on the radio relay link (Fig. 7). To perform this action, let’s find the intersection points of $BER = 10^{-6}$ (BER threshold in most modern duplex transmission systems) on the graph BER and E_b/N_0 for different MCS modes.

The thick solid lines in Fig. 7 represent the MCS modes at which the system will operate depending on the E_b/N_0 values.

Table 4 shows the E_b/N_0 (h^2) thresholds and dynamic range D_{EbN0} . Each MCS mode operates within these dynamic ranges. Let's display the energy transitions switching, in other words – dependence of maximum data rate from energy in radio relay link, in a sequential algorithm (Fig. 8).

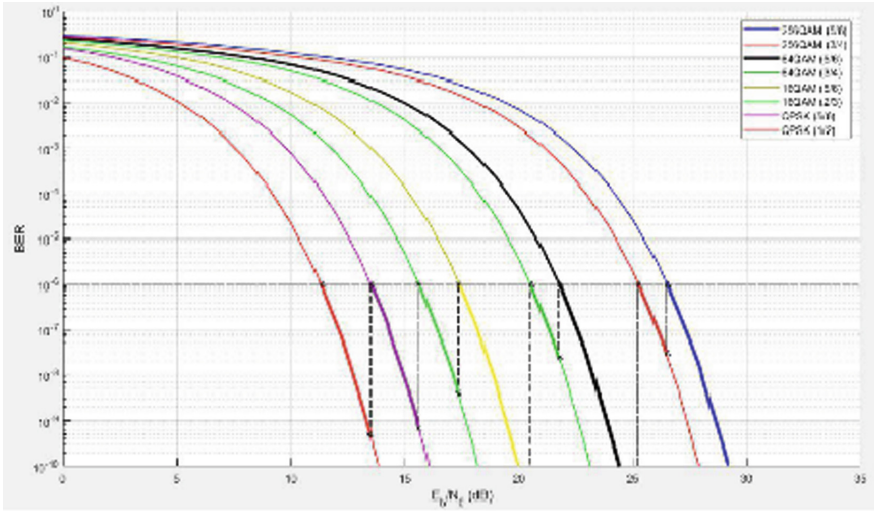


Fig. 7. Determining the threshold values of E_b/N_0 (h^2)

Table 4. E_b/N_0 dynamic ranges for MCS modes

MCS mode	D_{EbN0}
QAM-256 (5/6)	>26.52 dB
QAM-256 (3/4)	25.25–26.52 dB
QAM-64 (5/6)	21.76–25.25 dB
QAM-64 (3/4)	20.49–21.76 dB
QAM-16(5/6)	17.35–20.49 dB
QAM-16(2/3)	15.59–17.35 dB
QPSK (5/6)	13.48–15.59 dB
QPSK (1/2)	11.27–13.48 dB

For clarity let's construct the dependence of data rate on radio relay link from a distance between the receiver and transmitter. For this purpose, let's first express the parameter E_b/N_0 (h^2) as a function of the distance d (5) ... (8) (Fig. 9):


```

R(h2) := if h2 ≥ 26.52
    || 183.3
else if (h2 < 26.52) ∧ (h2 > 25.25)
    || 165
else if (h2 < 25.25) ∧ (h2 > 21.76)
    || 137.5
else if (h2 < 21.76) ∧ (h2 > 20.49)
    || 123.8
else if (h2 < 20.49) ∧ (h2 > 17.35)
    || 91.7
else if (h2 < 17.35) ∧ (h2 > 15.59)
    || 73.3
else if (h2 < 15.59) ∧ (h2 > 13.48)
    || 45.5
else if (h2 < 13.48) ∧ (h2 > 11.27)
    || 27.5
else
    || 0
    
```

Fig. 8. Data rates based on the energy conditions and selected MCS mode

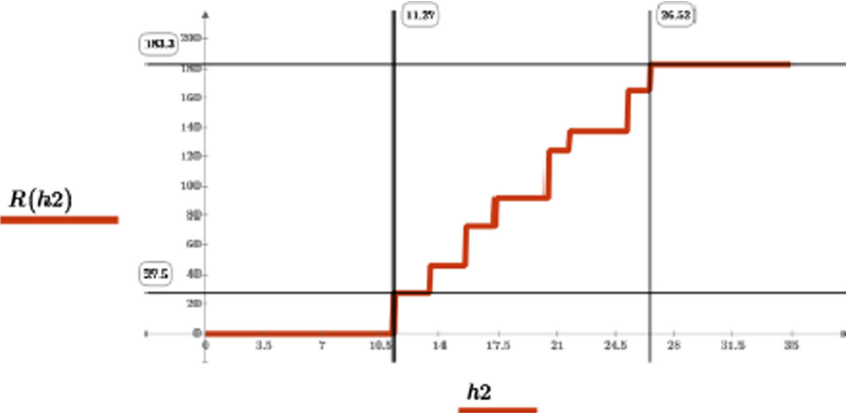


Fig. 9. Dependence of bitrate in radio relay link from signal energy that falls on a received 1 bit E_b , to the noise power spectral density N_0 using the adaptive MCS selection algorithm

$$EbN0(d) := SNR(d) + 10 \cdot \log\left(\frac{\Delta f}{R}\right) \tag{5}$$

$$EbN0(d) := P_{RX}(d) - (-120) + 10 \cdot \log\left(\frac{\Delta f}{R}\right) \tag{6}$$

$$EbN0(d) := P_{TX} + G_{TX} - L_F - L_{FS}(d) - L_O - L_G(d) - L_A + G_{RX} - (-120) + 10 \cdot \log\left(\frac{\Delta f}{R}\right) \quad (7)$$

$$EbN0(d) := P_{TX} + G_{TX} - L_F - (32.4 + 20 \cdot \log(f) + 20 \cdot \log(d)) - L_O - a \cdot d - L_A + G_{RX} - (-120) + 10 \cdot \log\left(\frac{\Delta f}{R}\right) \quad (8)$$

The graphical representation of dependency the $E_b/N_0(d)$ for the QAM-256 (5/6) mode is shown on Fig. 10.

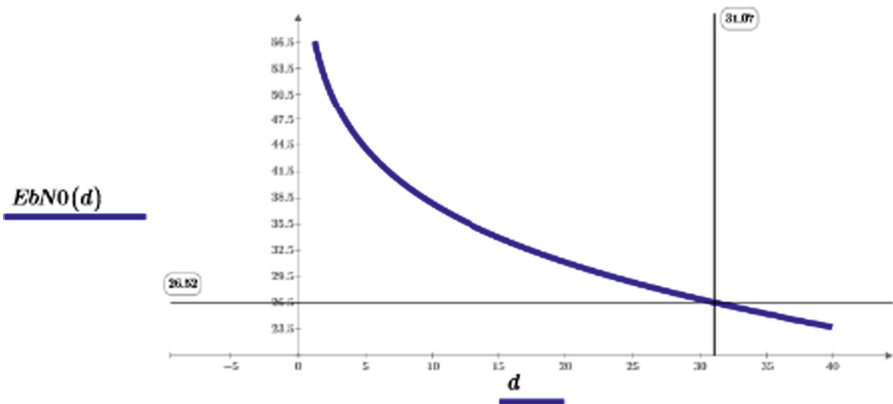


Fig. 10. The graphical dependency of signal energy ratio that falls on a received message 1 bit E_b , to the noise power spectral density N_0 , from the distance d between the receiver and the transmitter (the length of radio relay link)

As noted earlier, switching from the QAM-256 (5/6) mode to the QAM-256 (3/4) mode carried out when reaching the threshold $h^2 = 26.52$ dB (Table 4). Figure 10 shows that the distance 31.07 km corresponds to this h^2 value. Thus, if the planned radio relay link is longer than 31 km, then at a given hardware configuration it's not possible to achieve the maximum possible throughput with QAM-256 (5/6) mode as the BER value in the QAM-256 (5/6) mode is too high (over 10^{-6}), so switching at least to the QAM-256 (3/4) mode is mandatory.

Similarly, we define the threshold d , km, for all possible MCS modes of adaptive MCS operation (Table 5).

Table 5. Thresholds d for MCS modes in AWGN channel

MCS mode	Data rate R , mbps	d , km
QAM-256 (5/6)	183.3	31.07
QAM-256 (3/4)	165	36.33
QAM-64 (5/6)	137.5	52.29
QAM-64 (3/4)	123.8	59.96
QAM-16 (5/6)	91.7	83.04
QAM-16 (2/3)	73.3	99.56
QPSK (5/6)	45.5	128.06
QPSK (1/2)	27.5	161.83

Figure 11 shows the dependence of the maximum data rate and radio relay link length.

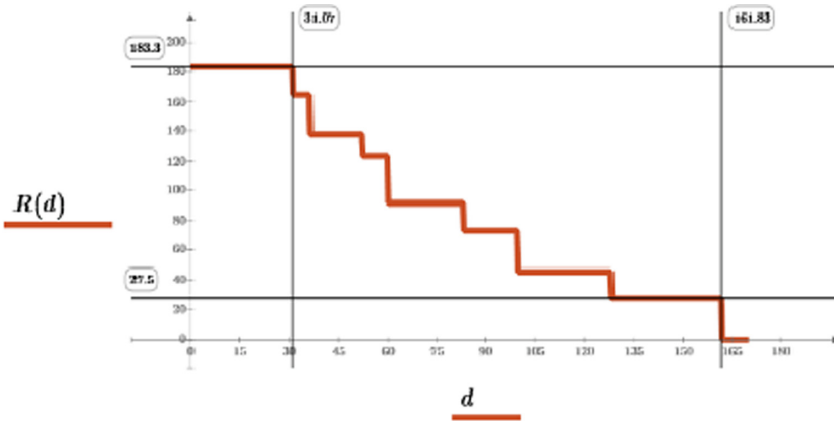


Fig. 11. The dependence of the maximum data rate and radio relay link length between the receiver and transmitter

It should be noted that the calculations above are performed for radio relay link that works in almost ideal conditions:

- Noise is additive white Gaussian noise;
- Signal fading is absent due to no multipath propagation on the line, attenuation in rain, etc.;
- Perfect antennas alignment performed;
- Line of sight and without interference.

In addition, in accordance to theoretical calculation, the radio relay link can work in the lowest MCS mode QPSK (1/2) at distances up to 160 km. But on practice it is almost impossible because of earth curvature, a very high probability of getting noise in the first Fresnel zone, the curvature of transmitting antenna main lobe due to changing

atmosphere gradient density and composition, and other effects that occur with long radio relay links.

4.2 Comparative Analysis of the Bit Rate Depending on the Distance Between the Transmitter and Receiver for Non-ideal Channel Conditions and Proposals to Improve Noise Immunity in Radio Relay Lines

Let's assess how bit rate depends on the distance between the receiver and transmitter in radio relay link with non-ideal conditions by introducing the additional fading to the signal. The most probable are the following cases:

1. The impact of **multipath**;
2. **Semi** open line of sight;
3. Heavy **rain** at the link path.

Let's estimate the influence value through multipath propagation which is the most common case in urban areas. For this purpose the Matlab mathematical model helps to build the graph of bit error BER on E_b/N_0 for QAM-256 (5/6) and QAM-64 (5/6) modes for the channel with white Gaussian noise and Rice channel. Thus, when considering multipath in radio relay communication link, then trying to choose the parameters of Rice channel as close as possible to real conditions, so it is assumed that most of the energy (about 90–95%) is concentrated in the main beam. Let assume energy losses are 7.5%. Then the K-factor (the ratio of energy concentrated in the main beam to the energy that goes to minority rays) in Rice model for the channel will be equal to $92.5/7.5 \approx 12.3$. Let's take a diversity factor equal to 1 so far as spatial antennas diversity in this case is not used. So, the following graphs received (Fig. 12).

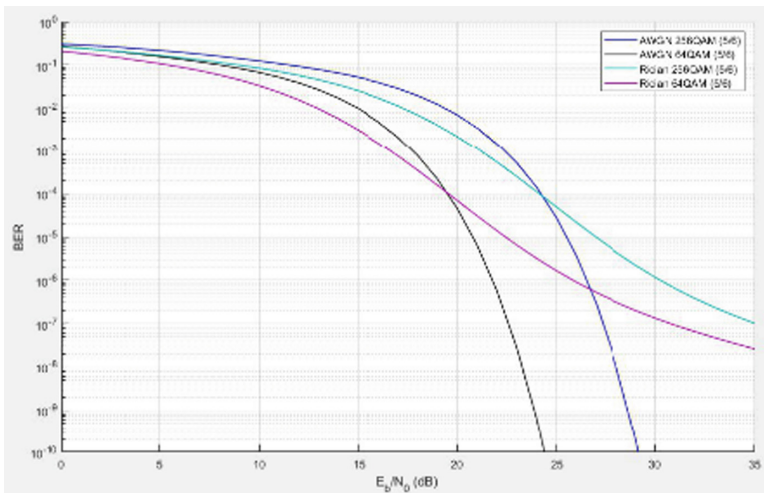


Fig. 12. Graphic representation of BER depends on E_b/N_0 for AWGN and Rice channels in different MCS modes

As can be seen on Fig. 12, the multipath effects the shift thresholds E_b/N_0 at $BER = 10^{-6}$, which will be switching to lower MCS modes around 3.5–4 dB compared to the AWGN channel. To simplify further calculations, let's present the SNR value lower to $L_{AI} = 4$ dB due to multipath fading as an additional fading.

Then the analytical dependence $E_b/N_0(d)$ will be as following:

$$EbN0(d) := P_{TX} + G_{TX} - L_F - (32.4 + 20 \cdot \log(f) + 20 \cdot \log(d)) - L_O - a \cdot d - L_A - L_{AI} + G_{RX} - (-120) + 10 \cdot \log\left(\frac{Af}{R}\right) \quad (9)$$

Thus, the dependency E_b/N_0 from d falls by 4 dB (Fig. 13).

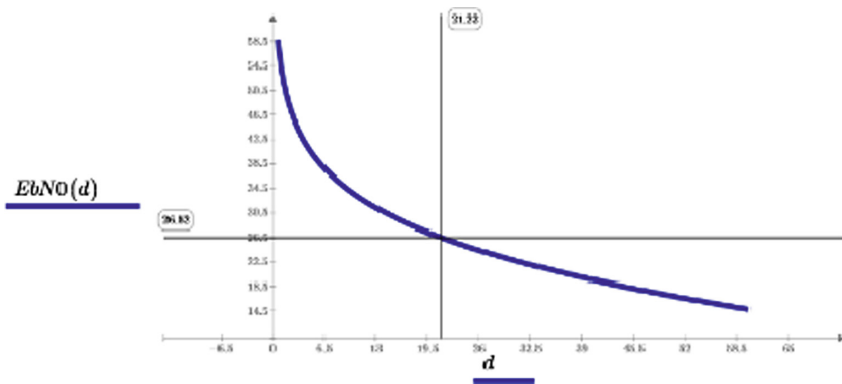


Fig. 13. Graphical dependency of SNR ratio falls on a received 1 bit energy E_b , to the noise power spectral density N_0 , from the distance between the receiver and transmitter when using the QAM-256 (5/6) mode taking into account the multipath phenomenon

The graph on Fig. 13 shows that multipath fading cause switching from the QAM-256 (5/6) mode to QAM-256 (3/4) mode closer to transmitter: it will be performed not on the distance of 31.07 km from the transmitter, but at a distance of 21.22 km, what clearly demonstrates the effect of multipath propagation. A similar situation exists for all MCS thresholds (Table 6).

Table 6. Thresholds d for MCS modes in Rice channel with multipath propagation and fading

MCS mode	Data rate R , mbps	d , km
QAM-256 (5/6)	183.3	21.22
QAM-256 (3/4)	165	25.09
QAM-64 (5/6)	137.5	37.25
QAM-64 (3/4)	123.8	43.28
QAM-16 (5/6)	91.7	62.05
QAM-16 (2/3)	73.3	75.97
QPSK (5/6)	45.5	100.72
QPSK (1/2)	27.5	130.95

So, in order to improve a noise immunity of the considered radio relay channel, it's suggested the use the adaptive MCS modes switch instead of the static switch. As the original set of MCS modes the DVB-T2 set is taken for analysis. The performed analysis shown the MCS modes redundancy, thus the MCS modes reduction number is recommended (Table 3).

The analysis of the proposed reduced set of MCS modes with influence of multipath effects shown that the proposed adaptive MCS modes selection allows a large set of equidistant throughput capacity values (from 27.5 to 183.3 Mbps), and a rather wide dynamic range E_b/N_0 that guarantees high system throughput under normal conditions of signal propagation and high noise immunity in the presence of interference, multipath and so on.

5 Conclusion

The peculiarities and features of high data rate duplex communication system (HDRDCS) is analyzed in the paper. It is determined that the main ways to ensure high noise immunity of communication systems are: increasing SNR by increasing the transmitter power, rational choice of MCS with high index modulation and noise immune codes, and others.

The technological features of the DVB-T2 television broadcasting standard are analyzed, such as encoding, modulation, interleaving (bit, symbol, time and frequency), OFDM, and others. The expediency of borrowing the tools and techniques that used in the DVB-T2 standard is considered for a HDRDCS to improve noise immunity.

The characteristics of the physical nature and especially their manifestations, that list all the factors that influence the transmission signal in a radio link, are researched. It was found that losses may appear in communication channel due to propagation in free space, refraction, diffraction, reflection, scattering, and absorption, attenuation in the rain, and due to other reasons. The fading effects from multipath propagation are considered in details. The choice of channel fading models is justified for further calculations. Considered two channels types with multipath propagation and fading: Rice and Rayleigh channels, including the finally selected Rice channel model.

The feasibility analysis of DVB-T2 standard features for HDRDCS analyzed to improve noise immunity.

The use of adaptive MCS modes selection algorithm is proposed to improve noise immunity of radio relay link instead of static MCS modes selection algorithm. As the original set of MCS modes the DVB-T2 MCS list is chosen. The analysis showed the MCS modes redundancy using this set, thus the reduction of the MCS modes number is proposed.

For the reduced MCS modes set the thresholds SNR (h^2) are calculated without fading and with fading using the Rice channel model. The thresholds show the SNR points for MCS modes switch to lower MCS modes in case of fading in the channel. The dependence of system throughput of radio relay link as a function of distance d is calculated. The calculation result allows simplifying the radio relay systems and algorithms.

The dependency of radio relay link bitrate as a function of distance between the receiver and transmitter for non-ideal conditions are considered by introducing additional fading in the signal link. The multipath propagation scenario is taken as non-ideal channel conditions. The results shown that the proposed adaptive MCS modes selection algorithm allows a large set of equidistant capacity values (from 27.5 to 183.3 Mbps) and a rather wide dynamic range E_b/N_0 , which guarantees high throughput system under normal channel conditions, and high noise immunity in the presence of multipath, noise, rain and so on.

The analytical results obtained in this research can be used for the algorithms of radio relay equipment functionality development in company vendors, for planning and construction duplex transmission systems by engineers and telecom operators, and for further research in the field of radio communications.

References

1. Hepko, I.A., Oleynik, V.F., Gull, Y.D., Bondarenko, A.V.: Modern wireless networks: state and prospects of development. In: Oleynik, V.F. (ed.) ECMO, Kyiv, 672 p. (2009)
2. Digital Video Broadcasting (DVB); Implementation guidelines for a second generation digital terrestrial television broadcasting system (DVB-T2). - Draft ETSI TS 102 831 V1.2.1 (2012)
3. Uryvsky, L., Solyanikova, V.: Features of multiservice streams transmission in DVB-T2 standard. In: IX International Scientific Conference of Students and Graduate Students "Prospects of Development of Information and Telecommunication Technologies and Systems", PRITS 2017: Abstracts of the Conference –K.: CPI them. Igor Sikorsky (2017)
4. Frame structure channel coding and modulation for a second generation digital terrestrial television broadcasting system (DVB-T2). - Draft ETSI EN 302 755 V1.4.1, February 2015
5. Narytnyk, T., Volkov, V., Utkin, J.: Radio relay and troposphere transmission systems. K.: Basis, 696 p. (2008)
6. Uryvsky, L., Osypchuk, S.: The analytical description of regular LDPC codes correcting ability. *Transp. Telecommun.* **15**(3), 177–184 (2014). ISSN 14076160
7. Uryvsky, L., Osypchuk, S.: OFDM signal research with varied subcarriers number. *Transp. Telecommun.* **17**(3), 192–201 (2016). ISSN 14076179
8. Fading Channels [electronic resource] MathWorks. <https://www.mathworks.com/help/comm/ug/fading-channels.html#a1071598663b1>
9. Rayleigh distribution [electronic resource] Wikipedia. https://en.wikipedia.org/wiki/Rayleigh_distribution
10. Rice distribution [electronic resource] Wikipedia. https://en.wikipedia.org/wiki/Rice_distribution



Time and Synchronization in Telecoms

N. L. Biriukov^(✉)  and N. R. Triska^(✉) 

National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Peremoga Avenue 37, Kyiv 03056, Ukraine
{nlbir2, ntriska}@ukr.net

Abstract. The chapter deals with the main principles and trends in the field of synchronization in telecommunications. The presented approach is based on the respective ITU-T research and standardization activities as well as the authors’ own studies and academic subjects in Institute of Telecommunication Systems of “Igor Sikorsky Kyiv Polytechnic Institute”. A significant part of the results were obtained by the authors during their work at the Ukrainian Research Institute of Communications – UNDIZ (Kyiv, Ukraine).

The fundamental definitions of network synchronization and transmission modes as well as the actual requirements for frequency synchronization (FS), phase (PS) and time of day (ToD) synchronization and their metrics are analyzed. The relationships and limits of FS and PS characteristics are proposed based on the FS and PS conceptions indivisibility. The two-way precise time protocol (PTP) is analyzed subject to time delay random component and frequency deviation. The analysis of both well-known and new metrics evaluating the synchronization accuracy was performed. The PTP time stamps exchange was studied using Maximum Likelihood method and minimal time interval errors *mTIE*.

Results and challenges for further study are summarized in the Conclusions.

Keywords: Synchronization · Frequency · Phase · Slips · Synchronous mode · Plesiochronous mode · Synchronous transmission way · Asynchronous transmission way · Metrics · Time interval · TIE · MTIE · MTIE · MATIE · MAFE · PTP · PDV

1 Introduction

Synchronization¹ is a fundamental phenomenon of nature, which in the broadest sense means any alignment of processes in time [1–5]. The phenomenon of synchronization covers the rhythms of both biological and cosmic worlds. Environments – “intermediaries” that provide capture, tracking processes in time, can be represented by fields of different nature: mechanical (including acoustic), electromagnetic, gravitational, etc. These fields represent the tool of influence – the environment for interaction of objects of diverse nature, which in certain circumstances creates the phenomenon of synchronization – the harmonization of the rhythms of living organisms, flickering fireflies

¹ The term “synchronization” comes from the Greek word *synchronos*, which means “simultaneously”.

or the movement of planets, biological and sociological rhythms, from applause to revolutionary sentiments, from the transmission and reception of information in communication systems to the dynamic synchronization of programmed networks [1–5]. According to the definition of academician V. Vernadsky, “... the desire for synchronization is one of the manifestations of the self-organization of nature.”²

The principles and approaches of synchronization techniques have evolved in many technical applications, as well as in the natural and human sciences. In a generalized form, the principles of synchronization are common to several general scientific, interdisciplinary branches, in particular, theories of oscillation, bifurcation, catastrophes and synergetic. These scientific directions and synchronization share the fundamental principles, namely: the presence of systems capable of generating the rhythms of signals of physical, biological, physiological or social nature; the presence of “soft”, indirect, non-rigid communication, which usually has a wave nature; and, finally, the ability of (dynamic) system to acquire a new quality (sometimes by a jump) due to a slight change in parameters.

Synchronization is not only a process of coordinating events, but also a research method, for example, a synchronic method in historical science – comparing the events of one time that took place in different countries and among different peoples³, as well as a method for studying the stability of networks in relation to so-called synchronization loops – cycles [2, 3, 26]. Methods for determining longitude in navigation are an example of a practical comparison (alignment) in time when determining the longitude of objects.

From the point of view of theoretical and practical application, synchronization in telecommunications covers all the above-mentioned areas and has acquired a wide, multivalued meaning, which requires more detailed consideration.

2 Basic Terms, Definitions and Topical Tasks

The concept of synchronization is ambiguous and, depending on the application, may have significant differences.

Firstly, it should be noted that tasks of achieving synchronization directly in the environment of transmission and switching systems and tasks of frequency, phase and time synchronization in telecommunication networks (TN) can be separated. The tasks of the first group arose simultaneously with the development and implementation of telecommunication systems and were solved during the development the system itself. Therefore, this class of tasks was almost never distinguished or emphasized, but it was always carefully developed in the design, debugging and operation of systems [1–3], for example, telegraph systems, PCM, PDH, SDH, ATM, Ethernet (BASE), Ethernet-TP, IP-MPLS, etc. Each of these systems at the physical and channel level provides frequency synchronization – f (clock frequency, bit) and phase – φ synchronization directly in the process of transmitting information in real time. Phase synchronization at

² Quoted from [4], in the authors’ translation from Russian.

³ “Encyclopedia” F.A. Brokhaus and I.A Efron in 86 volumes: St. Petersburg, 1890–1907.

the channel level of the OSI⁴ model is called cyclic or packet synchronization, depending on the way of transmitting information fragments. Detailed features of telecommunication network operation modes and transmission methods based on OSI model level and in terms of synchronization will be discussed below in Subsect. 3.2.

The tasks of network synchronization, that is, special, dedicated tools using the transport network to provide frequency, phase and time as services for its own needs as well as for other users, differ from the above-mentioned tasks related to the telecommunication systems development.

The synchronization mode – whether inside the system or the network – can be provided only in two ways: forced (master-slave mode [G.810]⁵, despotic [3]) or mutual. A mutual method is used in special cases, for example, to reserve multiple primary sources of the same quality, or in local computer networks.

During the so-called “classic” period of TN synchronization (80–90th years of the XX century) the basic principles of forced clock synchronization (timing network synchronization – TNS) in the environment with synchronous transmission and connection-oriented systems were developed. In addition, standards for the degree of accuracy and stability of signals and equipment and methods for checking compliance with standards were developed. These fundamentals, well-worked out in practice, do not lose relevance today, serving as the basis for developing solutions for synchronizing of the new generation TN [6–14].

The main distinction of the modern period of network synchronization is to provide frequency synchronization (FS), phase synchronization (PS) and time scales alignment (ToD) in conditions of asynchronous transmission – in a packet environment. This means that the phase in different geographically remote locations of the network must be the same [G.8260]. In other words, events, for example, certain significant moments (Fig. 1) should occur at the same time, and the time scales indications (time of day – ToD) must coincide.

The current challenges (and, accordingly, current ITU-T and other standardization studies) in the area of synchronization are related to the creation of the Internet of Things (IoT) infrastructure, 4G and 5G mobile networks (IMT-2020) and to support of telecommunication infrastructure transition from NGN to software-oriented networks SDN/NFV⁶ [15, 16]. In this context, methods and tools for evaluating the quality parameters of packet transmission should be considered, taking into account the possibility of providing an appropriate accuracy level of frequency and phase parameters.

⁴ OSI – Open Systems Interconnection (OSI/IEC 7498-1:1994).

⁵ Hereinafter, the references to the ITU-T Series numbers of the G series are given in square brackets. The current versions of all ITU-T Recommendations are available on the official website: <https://www.itu.int/>.

⁶ SDN/NFV – Software Defined Network/Network Functions Virtualization.

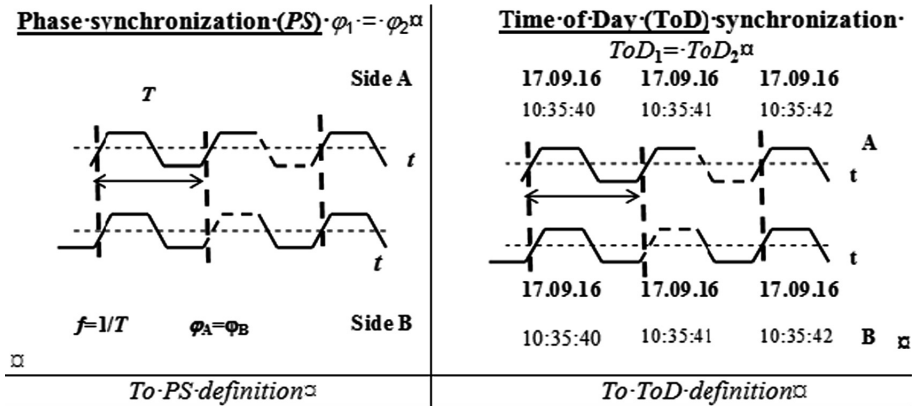


Fig. 1. Phase (PS) and time (ToD) synchronization in the telecommunication network

3 Timing Signal Model and Synchronization Networks Modes

3.1 Timing Signal Model

The model for measurement results estimating, slip frequency calculating and metrics analysis is based on the measurement of the deviation, the phase difference $TE(t)$ (Time Error) between reference (Ref) and working (UT – under test) timing signals (see Fig. 2).

Time Interval Error – $TIE(t; \tau)$ is evaluated by the uniform $TE(t)$ samples $\{x_i = x(i\tau_0), i = 1, 2, \dots, N\}$ of the difference between the significant moments of the measured timing signal $T(t)_{UT}$ and the reference signal $T(t)_{ref}$ within the measurement period $T = (N-1)\tau_0$, where τ_0 is the sampling interval and N is a number of equally spaced samples of $x(t)$.

Assuming that the reference clock instability is negligible as compared with the clock under test, the general model of phase offset $x(t)$ may be presented as [1–3]:

$$x(t) = x_0 \pm y_0 \cdot t \pm Dt^2/2 + \varepsilon(t) + j(t), \tag{1}$$

where: $x(t)$ – the current phase offset $TE(t)$, x_0 – initial phase offset; $y_0 = |f_0 - f_n|/f_n$ – initial frequency offset regarding to the nominal frequency f_n , D – frequency drift, $\varepsilon(t)$ – random phase noises of clock sources and PLL⁷ devices, $j(t)$ – non-random, algorithmic (for, example, stuffing) and other components of the spurious phase modulation.

It follows from model (1) that in the presence of at least only the initial frequency offset, the deviation of the significant moments of the signals from two

⁷ PLL – Phase-Locked Loop.

non-synchronized generators will increase – there is a “problem of two generators”⁸ when the recording of an information signal is carried out with one frequency while the reading is carried out with another one [1, 2, 17].

$$\begin{aligned} TE(t) &= x(t) = T(t)_{UT} - T(t)_{ref} \\ TIE(t, \tau) &= x(t + \tau) - x(t) \end{aligned} \tag{2}$$

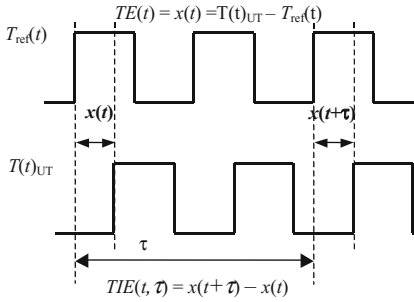


Fig. 2. Definitions of $TE(t)$ and $TIE(t, \tau)$; $T(t)_{UT}$ – timing signal under test, $T(t)_{ref}$ – reference timing signal; $x(t)$, $x(t + \tau)$ – the difference between significant instants of a signal under test and a reference signal for t and $t + \tau$, respectively

Based on the known relationship between frequency and phase $y(t) = dx(t)/dt$, we obtain a model for the normalized frequency:

$$y(t) = y_0 \pm Dt + \varepsilon(t)' + j(t)', \tag{3}$$

where all symbols correspond to the explanation in (1), and $\varepsilon(t)', j(t)'$ are the derivatives of $\varepsilon(t)$ and $j(t)$, respectively.

In the general case, the deviation of time interval $TE(t)$ is a result of the influence of systematic components (initial frequency offset y_0 and frequency drift D , algorithmic phase distortions) and random components (phase or frequency modulation of clock signals by white or flicker noise).

In addition to the initial setting error, the frequency offset may be caused by the Doppler effect. If objects move at a constant rate, then the Doppler effect is manifested in a constant frequency offset and so reduces to the case under consideration. If the objects were moving with acceleration, there would be a drift of frequency due to the Doppler effect.

Thus, $TE(t)$ is a non-stationary, time-dependent, random variable. In cases of conformity assessment, the measurement uncertainty and possible calculation errors should also be considered.

Models (1), (3) change their content depending on the network synchronization mode. So, if we estimate the phase shift in the plesiochronous connection (in the

⁸ See Fig. 4a below.

plesiochronous network – PN) then all components of formulae (1), (3) should be considered while in the synchronous connection only the phase noise $\varepsilon(t)$ and algorithmic phase distortions $j(t)$ will be relevant. In PN the most important component is $y_0 \cdot t$, but frequency shift D should be considered only after several days of observation.

3.2 The Main Modes of Transport Network Synchronization

As noted above, the approaches and technical means for solving synchronization tasks depend on the mode in which the network operates on the physical level of the OSI model as well as on the way of information fragments transmission at the channel level [1, 2, 6, 7, 13]. The main variants of TN synchronization modes and transmission methods are summarized in Fig. 3.

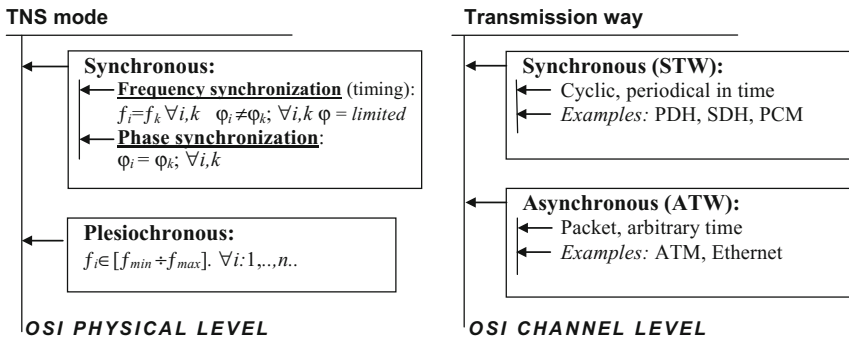


Fig. 3. Place of transport technologies (level 1 and level 2) in the OSI model

At the physical level TN synchronization during real-time transmission can be provided in *synchronous* or *plesiochronous* modes. The synchronous mode is considered here as the forced synchronization of all TN elements from a reference source (or multiple sources) of high quality. In the classical timing network synchronization (TNS) the frequencies are on average the same ($f_i = f_k \forall i, k$) and the phases are arbitrary – not identical, but with a limited shift. For phase synchronization (Fig. 1), besides the same frequency, the TNS network should also provide the same phase $\varphi_i = \varphi_k$ for all network elements. The plesiochronous mode does not provide the forced synchronization, so all generators in the network operate in free-run mode, but the frequency deviation is limited: $f_i \in [f_{min} \div f_{max}] \forall i: 1, \dots, n.$

Depending on the way of information fragments transmission at the channel level, synchronous and asynchronous transmission ways are distinguished (STW and ATW, respectively). STW is characterized by a cyclic, periodic signal structure, while the ATW involves the division of the digital stream into fragments of different lengths (packets) with subsequent packet transfer in a relative time scale [1–3]. As regard TNS providing, STW systems can operate in both synchronous and plesiochronous modes (depending on transmission and switching technologies and specific design decisions). ATW networks are, by definition, plesiochronous, but the need for precise

frequency/phase/time synchronization forces developers to improve ATW systems using traditional TNS approaches adopted for packet environment [7–14].

In any case, the network synchronization characteristics have a significant impact on the quality of TN operation, especially for real-time telecommunication services. At each stage of telecommunications development, network synchronization parameters determine the achievable level of TN operation performance and the most effective ways to implement and distribute synchronization services. From the point of view of synchronization, STW and ATW networks have both advantages and disadvantages, especially in the case when the level of real-time operation quality depends on the change in packet delay (packet delay variation – PDV).

Subject to the current trends of transition to fully packet transmission and switching principles and the corresponding changes in TN infrastructure, the plesiochronous TNS mode (plesiochronous network – PN) is of particular interest as a more general case of network operation [18–20]⁹. Therefore, this mode will be analyzed below in the context of slips occurrence and prevention.

4 Slips

4.1 Definition

The main purpose of synchronization networks is to ensure the quality of information transmission, i.e. the correct (without errors) operation of the telecommunication network. The discrepancy of frequencies of the network elements clocks leads to phase deviation in the equipment, which, in turn, causes the slips in the transmission process. *Slip* is a typical error due to synchronization failures: loss or repetition of a digital stream fragment (bit, byte or frame, packet or ATM cell) [1–3, 17–19].

4.2 Slips Frequency/Period Estimation

The discrepancy of frequencies in each network element – the so-called “problem of two generators” – has certain restrictions on the phase difference associated with the capacity of elastic/buffer memory (EM) used in transmission and switching equipment. An example of the physical and logical structure of the EM is shown in Fig. 4 [18–20].

The clock sources have limited accuracy and stability, and therefore the actual values of clock frequencies during writing and reading differ from the nominal ones. As a result, in the rewriting device (for example, in EM) there is a difference of writing and reading clock frequencies¹⁰ denoted in Fig. 4 as y_1 and y_2 (in normalized form). The current values $y_1(t)$ and $y_2(t)$ are in the normalized interval $[-1, +1] \cdot y_0$. Note that the EM operation is shown in Fig. 4 in real time. The model of EM operation in relative

⁹ It should be noted that, taking into account G.701 and G.822 ITU-T Recommendations, the quasi-synchronous and asynchronous synchronization network modes described in G.810 can be considered as special cases of the plesiochronous network mode.

¹⁰ Here and below it is assumed that the nominal values and the permissible offsets of writing and reading frequencies are equal.

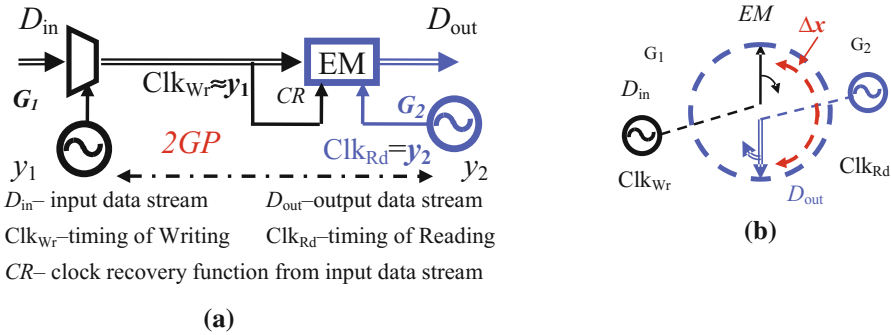


Fig. 4. (a) Elastic/Buffer memory (EM) in the digital path of one direction in PN for decision of “problem of two generators” G_1 and G_2 (2GP) (b) Logical EM model

time, for example, in the case of waiting in the queue, has not fundamental differences from the model in Fig. 4 [18–20].

The general practice of slip time estimating is based on the assumption of maximum frequency deviation, that is, the frequencies have the extreme opposite values at an acceptable interval. Further, based on the permissible (limited) phase offset x_{lim} of two signals in (1) and neglecting other components, we can define the slip time as:

$$t_{sl} \geq x_{lim} / 2 y_0 \tag{4}$$

It follows from (4) that even if the initial inaccuracy does not increase, a slip occurs inevitably. The obtained estimation indicates the shortest period of slips. E.g., for typical values for a digital exchanges $x_{lim} = 125 \mu s$ and $y_0 \approx 10^{-11}$, the time of possible slip is 72 days.

It should be noted that if the frequency difference probability is estimated only by the distribution at the two boundary points, then, under the condition of the same probability of distribution at the boundaries of the permissible interval, the probability of identical frequencies or their displacement on the opposite boundaries is $1/2$. In other words, the extreme deviation, equal to the value of the total permissible deviation interval, is achieved with a probability of 0.5.

In the general case, frequencies can have any values within the tolerance. This suggests that the actual ratios for slip period estimation can vary significantly in the larger direction. The most probable assumption is the normal frequency distribution within a normalized interval with an average value of y_0 and a maximum deviation of $\pm 3\sigma$ on $[-1, +1]$. Investigations by analytical methods and probabilistic modeling [18–21] showed that, under the normal distribution of the frequency values in permissible limits, the interval between slips increases at least twice with a confidence probability of 0.95. So, for the aforementioned values $x_{lim} = 125 \mu s$ and $y_0 \approx 10^{-11}$, the slip period estimation is almost 118 days (against 72 days) with a 99% probability.

The analysis of formula (4) confirms the importance of maintaining a higher signal accuracy at the network nodes, since when the accuracy degrades the slip frequency increases inversely proportional to the degree of accuracy. Permissible phase offset x_{lim}

is determined by the features of the reception method, for example, for bit-wise reception $x_{lim} \leq 0.45$ per unit interval.

It is of practical interest to estimate the frequency of slippage, the magnitude of the delay and its dispersion in the circuit of successively connected network elements that contain buffer memory devices. For example, modeling a hypothetical reference circuit [G.822] shows that the average value of the divergence of the write and read frequencies increases strictly in proportion to the number of series-connected EM devices, while the average time between slippage decreases approximately exponentially [34].

4.3 The Slip Preventing Methods

Most of the traditional synchronization tasks in telecommunications are aimed at preventing and/or minimizing of slips. Such methods can be conventionally divided into *non-network* methods (in PM using own resources of transmission systems) and *network* methods (the creation of a separate medium of one frequency based on the distribution network from one high-precision source).

Non-network methods of slip preventing include:

- rate alignment during the processing of plesiochronous signals (in PDH and SDH transmission systems);
- clock tracking, synchronization in the direction of transmission; it can be permanent connections (PDH and SDH line paths, loops at the digital exchange substation, synchronous Ethernet (SyncE)) or temporary connections (“tunnels”, channel emulation in ATW systems);
- use of the EM for smoothing of phase differences (in the equipment of transmission and switching systems based on STW and ATW);
- clock retiming increasing the timing signal quality (in the equipment of transmission and switching systems based on STW and ATW).

Non-network methods, except synchronization in directions, solve the slip preventing problem by limiting the duration of continuous transmission, i.e. “cutting” the continuous digital stream to the alignment parts or the packets, the duration of which does not exceed the slip period. The method of constraining of continuous transmission depends on the transmission method used at the channel level – STW or ATW.

In the case of synchronous transmission of information fragments in real time (STW), the direct and reciprocal alignment procedures are performed in a practically continuous stream (for example, the alignment of tributary signals in PDH and SDH transmission systems).

In the case of asynchronous transmission (ATW), the digital stream is pre-split into fragments of different lengths (packets), and packet transmission is performed in a relative time scale that is acceptable for transmitting data which is not critical to delays and/or time dispersions. As a result, the duration of the continuous transmission session is less than the expected slippage period. However, the scatter of packet sending and receiving times and the possible occurrence of queues during recording and reading lead to an appropriate delay variation. For some services the requirements to this delay variation are twice as more stringent than the delay value itself [18–20].

The method of synchronization in directions (paths) was a precursor to creating a synchronization network that removes the restrictions on location or temporality of the connections.

Network method of slip preventing is based on the timing network synchronization (TNS) concept, which effectively solves the slip problem by maintaining high accuracy and stability of the clock in all network. This is a forced method of maintaining a roughly equal mean frequency with a limited phase offset. This method is widely used both in “classic” STW-based networks and in modern ATW-based networks (for example, SyncE).

5 The Comparison of Synchronous and Plesiochronous Modes of Transport Network Synchronization

5.1 A Plesiochronous Network: Advantages and Disadvantages

Plesiochronous network (PN) is a more common and generalized way of interconnection of telecommunication networks, especially at the international level. As already mentioned in 4.3, PN has a wide range of means available to ensure the correct operation of telecommunication systems providing the carrier class transmission quality [9]. This allows you to effectively manipulate non-network synchronization tools, including methods of local timing support. Promising transport technologies, such as optical transport networks (OTNs), Ethernet-TP, IP/MPLS, suppose the operation in PN environment. In addition, the further development of local timing techniques will be stimulated by increasing of accuracy, broadening of range and reducing of cost of precise frequency and time sources, as well as by the presence of many developed global navigation satellite systems.

One of the directions for further research in PN may be the assessment of quality indicators in hypothetical chains of network elements and, in particular, the accumulation of the absolute value of the delay and its dispersion.

5.2 A Synchronous Network: Advantages and Disadvantages

A fully synchronized network with forced synchronization requires the implementation of a special rather complex and expensive multi-level TNS infrastructure [22–25]. TNS networks were created in all developed countries of the world in the 1990s and formed the so-called “classical” stage of TNS, which allowed providing the highest quality level that can be achieved in telecommunications (carrier class TN) and which cannot be directly achieved in packet-only environment [9, 13, 14].

But TNS networks in their original form have a number of significant drawbacks. Firstly, the higher costs mentioned above, and, secondly, the high complexity of design and operation. The last reason is caused by the possible uncontrolled slips, even in a synchronized connection.

By definition, in a synchronous network the phase offset is defined only by the phase noise in model (1), but in practice the wander and jitter accumulation as well as loops or “shuttle” switches can lead to short-term or long-term failures with a random

period of duration, cause or primary source of which is very difficult to detect [26–28]. The TNS sustainability is influenced by errors in network planning and programming of its elements as well as by operational procedures in case of loss and switching of reference synchronization signals due to emergency situations or during rebuilding and developing the network [13–25].

For TNS sustainability, the graph of the network should consist of a synchronization tree. The total number of root trees D for the distribution of synchronization signals in the network with N nodes is $D \leq N^{N-2}$ [26–28]. With increasing of number of nodes (N) and branches (R) in the network graph, the total number of root trees and, respectively, the potential threats of the occurrence of synchronization loops (cycles) also increases: $R_{\text{threat}} = R - (N - 1)$, especially during the operational network restoration.

Therefore, when choosing a TNS mode and transmission methods in the network, various factors should be considered to obtain an optimal synchronization scheme for the provided services.

6 Metrics for Synchronization Quality Evaluation

The metrics for estimation of synchronization signals, equipment and connections quality have been developed in the initial “classical” period of digital telecommunication networks deployment in 80–90th years of the XX century. For example, to evaluate the quality of EM operation in PN, the Maximal Time Interval Error $MTIE(\tau)$ function was proposed. The methods for evaluation of the first high-precision cesium generators, for example, the Allan variance and its modification as a Time Deviation $TDEV(\tau)$, had a significant impact on the international standardization process. The sufficiency of these two metrics is confirmed by many years of practice and international standardization [G.810 – G.813, G.823 and others].

The difficulties in evaluation of the conformity of transmission quality in packet-based, asynchronous environment with the requirements of previous technologies, has stimulated the search for new metrics [G.8260] and the expansion of classical metrics family including the $MTIE(\tau)$ and/or their adaptation and new interpretation [30–33].

6.1 Features of $MTIE(\tau)$

The detailed study of general and extended features of $MTIE(\tau)$ was performed in [29–32]. The $MTIE(\tau)$ function is estimated by the following formula:

$$MTIE(n\tau_0) = \max_{1 \leq k \leq N-n} [x_{ppk}(\tau_k)], \quad n = 1, 2, \dots, N - 1 \quad (5)$$

where $x_{ppk}(\tau_k) = \left[\max_{k \leq i \leq k+n} x_i - \min_{k \leq i \leq k+n} x_i \right]$.

For further analysis we should consider the following $MTIE(\tau)$ features: (1) it directly estimates a phase offset $TE(\tau)$ and $TIE(t, \tau)$; (2) it is sensitive to all components of the timing signal model (1) except a constant component; (3) by definition, $MTIE$

gives an upper estimate of all other time metrics in equal observation intervals; (4) due to exclusion of the constant component x_0 we can estimate the delay variation neglecting the absolute value; (5) $MTIE(\tau)$ features can be easily expanded to estimate average or minimal phase offset in the given observation interval; (6) $MTIE(\tau)$ graph has a specific shape due to its concavity; (7) $MTIE(\tau)$ can be normalized and interpreted as a distribution function in time domain.

6.2 The Expansion of $MTIE(\tau)$ Family

The expansion of metrics based on $MTIE$ function was proposed in [30] and extended in [31]. The peculiarity of these functions is that they inherit the main features of $MTIE(\tau)$ function. Subject to the practical application for the $TE(t)$ flow evaluation and regulation, it is advisable to determine at least the minimum peak-to-peak phase offset $mTIE(\tau)$ and the maximum frequency value $\hat{Y}_{pp}(\tau)$ at observation time τ [30, 31]:

$$mTIE(n\tau_0) = \min_{1 \leq k \leq N-n} [x_{ppk}(\tau_k)], \quad n = 1, 2, \dots, N/2 - 1 \tag{6}$$

$$\hat{Y}_{ppk} = \max_{1 \leq k \leq N-n} [x_{ppk}(\tau_k)] / \tau_k \geq \bar{y}_k, \tag{7}$$

where $x_{ppk}(\tau_k)$ is defined in (5), $\bar{y}_k = (x_{k+1} - x_k) / \tau_k$, $\tau_k = t_{k+1} - t_k$.

Formula (7) can be considered as a derivative of $MTIE(\tau)$ and, thus, $\hat{Y}_{pp}(\tau)$ can be considered as the empirical density of the distribution of maximum frequency values at the observation time intervals.

By definition, $mTIE(\tau)$ (6) has inherent properties of $MTIE(\tau)$. In order to estimate the minimum deviations of time stamps in the exchange protocols, the definition area in (6) is limited to a half of the controlled measurement interval.

Similarly, the mean values of the phase or frequency over time intervals, as well as the maximum and minimum values in the function ensemble [31] are determined.

6.3 Practical Applications of Metrics

The Use of $MTIE(\tau)$, $mTIE(\tau)$ and $MATIE(\tau)$, $MAFE(\tau)$

The practical value of the $MTIE(\tau)$ function is that it estimates the phase deviation $x_{pp}(\tau)$ accumulated between the clock signals during the EM filling or emptying period $t_0 \leq t_0 \leq t_0 + \tau$ (Fig. 4) proportionally to this deviation. Conformity with $MTIE(\tau)$ norms guarantees that the EM capacity is not exceeded at a given time interval τ , that is, $\max[x_{pp}(\tau)] \leq x_{lim}$ and so a slip does not occur. The transmission process in the synchronous network is characterized by a practically constant delay and a limited phase offset, the random component is absorbed in the EM and is smoothed in PLL circuits. As follows from 4.2, in PN the $MTIE(\tau)$ allows estimating the possibility of a slip.

As noted above, the asynchronous transmission way has significant differences. The value and nature of the delay depends on the operation mode of switches, routers, splitters etc., as well as on the time of day. In this regard, the use of some new metrics is proposed, in particular, the Maximum Average Time Interval Error – $MATIE(\tau)$ and the Maximum Average Frequency Error– $MAFE(\tau)$ [31]:

$$MATIE(n\tau_0) \cong \max_{1 \leq k \leq N-2n+1} \frac{1}{n} \left| \sum_{i=k}^{n+k-1} (x_{i+n} - x_i) \right| \quad n = 1, 2, \dots, N/2 - 1 \quad (8)$$

$$MAFE(n\tau_0) = MATIE(n\tau_0)/n\tau_0 \quad (9)$$

Comparison of formulae (6), (7) and (8), (9) shows that they have a close approach but different interpretations. Moreover, as will be shown below, $mTIE(\tau)$ can be used to evaluate the two-way time stamps exchange and obtain data close to the Maximum Likelihood method [30, 32, 33].

The Comparison of $MTIE(\tau)$, $mTIE(\tau)$ and $MATIE(\tau)$, $MAFE(\tau)$

By definition, at the given time intervals τ $MTIE(\tau)$ gives an upper estimate of time metrics [31]: $MTIE(\tau) \geq \{TIE(\tau), TDEV(\tau), TIE_{\text{rms}}(\tau), MATIE(\tau)\}$, while $\hat{Y}_{pp}(\tau)$ gives an upper estimate of the dimensionless, normalized metrics of the Allan deviation family: $\hat{Y}_{pp}(\tau) \geq \{ADEV(\tau), MADEV(\tau), MAFE(\tau)\}$, $\hat{Y}_{pp}(\tau) \geq MAFE(\tau) \geq mTIE(\tau)/\tau$.

$MATIE(\tau)$ and $MAFE(\tau)$ estimate the averaged values of time and frequency deviations and are not sensitive to phase offset in model (1), a part of the information is lost and not all factors can be evaluated [31].

The analysis of measurement results and modeling shows a high level of correlation between the metrics under consideration. $\hat{Y}_{pp}(\tau)$ and $MAFE(\tau)$ functions with the Gaussian phase distribution are the most correlated having a correlation coefficient not less than 0.9–0.95. $MTIE(\tau)$ and $mTIE(\tau)$ are correlated with a coefficient of more than 0.9.

In the presence of frequency synchronization, $MTIE(\tau)$ and $mTIE(\tau)$ indicate the relative value of the random phase deviation observed over the observation intervals, neglecting the constant component x_0 in (1). Thus, in the case of phase and ToD synchronization the current phase error $\Delta\varphi(t)$ is not less than $TIE(t, \tau)$, or $\Delta\varphi(t) \geq x(t) = y_0 \cdot t$. In general, $\Delta\varphi(t)$ changes proportionally to $TE(t)$ and the normalized frequency value: $\Delta\varphi(t) \sim x(t) \sim y(t)$ [13]. Subject to the definitions and Figs. 1 and 2, the following sequence of synchronization parameters evaluation is logical: FS \leftrightarrow PS \rightarrow ToD or $f \leftrightarrow \varphi \rightarrow T$.

7 The Analysis of Two-Way Precise Time Protocol

The simplified general model of the two-way time stamps protocol (for example, PTP protocol) is presented in Fig. 5 [32, 33]. Figure 5a shows the cycles of two-steps sync messages exchange as an ideal model, without a packet delay variation (PDV). The delay time in each cycle is estimated by (10), and the shift of the time scale is estimated by (11).

$$\hat{d}^{(MS)} = \hat{d}^{(SM)} = \frac{1}{2} \cdot [(T_2^S - T_1^M) + (T_4^M - T_3^S)] \quad (10)$$

$$\hat{s}^{(MS)} = \frac{1}{2} \cdot [(T_2^S - T_1^M) - (T_4^M - T_3^S)], \quad (11)$$

where $\hat{d}^{(MS)}$, $\hat{d}^{(SM)}$ – estimates of one-way delay for transmission from Master (M) to Slave (S) Clock and from S to M, respectively; $\hat{s}^{(MS)}$ – estimates of time scale shift.

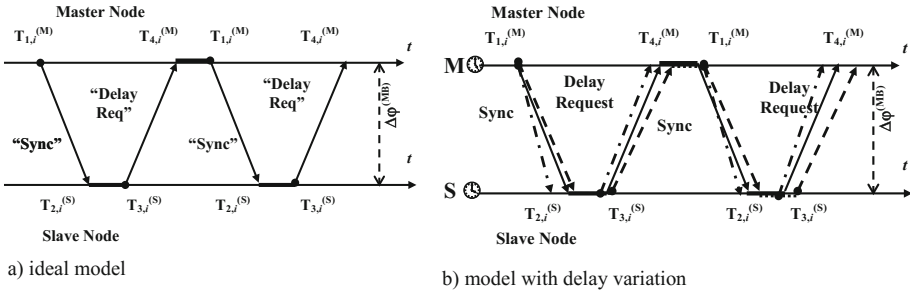


Fig. 5. Simplified general model of the two-way time stamps protocol

A more real model of time stamps exchange is shown in Fig. 5b, where the moments of any packet arrival is liable to a random delay and deviates around an ideal moment shown in Fig. 5a. In the both models it is supposed that Master and Slave Clocks are synchronized in frequency, i.e. the clocks have the same nominal frequency in each observation interval.

Let us denote the delay in the forward direction from M to S as $D(t)$ and in the opposite direction – as $U(t)$. Then (10) can be presented as:

$$D(t) = (T_2^S - T_1^M) = d^{MS} + \varepsilon(t)^{MS} + s^{MS} \quad (12)$$

$$U(t) = (T_4^M - T_3^S) = d^{SM} + \varepsilon(t)^{SM} - s^{SM}, \quad (13)$$

where $s^{MS} = s^{SM}$ is the phase difference between M and S clocks; d – a constant component of delay (it is assumed symmetric in both directions, i.e. $d^{MS} = d^{SM}$), $\varepsilon(t)$ – random components, independent and unequal for each direction, exponentially distributed in time, i.e. ε^{MS} and ε^{SM} are random parts of the time deviation functions $TE(t)^{MS}$ and $TE(t)^{SM}$, respectively. Thus, the received packets rate λ_r differs from transmission rate of sender λ by a random part $1/\varepsilon(t)$.

Substitution (12) and (13) in (10) and (11) gives:

$$\hat{d}^{(MS)} = \hat{d}^{(SM)} = \frac{1}{2} \cdot [D(t) + U(t)] \quad (14)$$

$$\hat{s}^{(MS)} = \frac{1}{2} \cdot [D(t) - U(t)] \quad (15)$$

When the both Master (M) and Slave (S) receive packets with propagation delay d provided that $\varepsilon^{MS} \neq \varepsilon^{SM}$, we have the following phase offset (time scale) estimations:

$$\hat{d}^{(MS)} = \hat{d}^{(SM)} = d_{ideal} + \frac{[\varepsilon(t)^{MS} + \varepsilon(t)^{SM}]}{2} \quad (16)$$

$$\hat{s}^{(MS)} = s_{ideal} + \frac{[\varepsilon(t)^{MS} - \varepsilon(t)^{SM}]}{2} \quad (17)$$

If protocol uses the multiple (N) cycles of message exchange, than there are two exponentially distributed data sets $\{D_i\}$ and $\{U_i\}$, $i = 1, 2, \dots, N$. So we can get time delay estimate by Maximum Likelihood method [30, 32]. For example, for exponentially distributed service time the delay estimate is $\hat{d} = (\min_{1 \leq i \leq N} D_i + \min_{1 \leq i \leq N} U_i)/2$ while for the Gaussian distribution we have $\hat{d} = (\bar{U} - \bar{D})/2$. The respective formulae for uniform distribution and a constant service time were obtained [30, 32]. By selecting random parts, we obtain an estimate of the random component of the delay time as $\hat{d}_s^{(MS)} = [\min_{1 \leq i \leq N} \{\varepsilon_i^{MS}\} + \min_{1 \leq i \leq N} \{\varepsilon_i^{SM}\}]/2$ and an estimate for the time scale as $\hat{s}_s^{(MS)} = [\min_{1 \leq i \leq N} \{\varepsilon_i^{MS}\} - \min_{1 \leq i \leq N} \{\varepsilon_i^{SM}\}]/2$.

The comparison of $mTIE(\tau)$ or $MTIE(\tau)$ with a Maximum Likelihood estimation (for example, in the assumption of the exponential distribution of $mTIE(\tau)$ or the determined $MTIE(\tau)$ service time) was performed. The simulation showed close estimation results for the mean, maximum and minimum phase deviations in a two-way exchange with a low mutual correlation ($<0,09$) [32].

Analysis of the simulation and field measurements results also showed the dependence of the $TE(t)$ variance on the value of frequency deviation. Frequency inaccuracy can be given by formula similar to (12) ÷ (15). This conclusion coincides with the conclusion of the PN analysis, which shows that the $TE(t)$ distribution function at the output of the EM is one-sided and has a constant component of the delay [18, 19]. The distributions of the $TE(t)$ in PM and the packet delays are not significantly different in shape. In the latter case, the significant issue is considerably greater values of packet delay and its deviation (PDV).

8 Conclusions

The current challenges in the area of synchronization are related to the creation of the Internet of Things (IoT) infrastructure, 4G and 5G mobile networks (IMT-2020) and to support of telecommunication infrastructure transition from NGN to software-oriented networks SDN/NFV. The study of the packet-based, asynchronous transmission features to ensure the compliance with frequency (f), phase (φ) and time (ToD) synchronization requirements also remains actual.

In the context of implementation of packet TN infrastructure, there is a need for adaptation and updating of traditional synchronization approaches and technical solutions (“classical” TNS) to provide FS/PS/ToD in the ATW-based plesiochronous network. Important steps along this path are the further study of synchronization processes in packet networks (in particular, the research of two-way time stamps (ToD) exchange via PTP protocol) and search of optimal evaluation means (metrics). One of the directions for further research may be the assessment of quality indicators in hypothetical chains of packet network elements and, in particular, the accumulation of the absolute value of the delay and its dispersion.

The study of Master (M) and Slave (S) clocks interaction in packet network has been performed based on the proposed two-way time synchronization model that includes random components. Some possible ways to estimate the random parts of time delay and time scale (phase) offset for two-way time stamps exchange were proposed.

The obtained results have shown that the metrics based on $MTIE(\tau)$, due to their features, can be effectively used for integrated evaluation of transmission performances in the packet network, in particular, to evaluate two-way time stamps (ToD) exchange via PTP protocol.

All models and metrics used for synchronization signals analysis are based on the measurement of the phase deviation $TE(t)$ followed by the time interval error $TIE(t; \tau)$ calculation, which in turn is a basis for calculation of the standardized metrics. The thorough research of the basic, well-known metric – the maximum time interval error $MTIE(\tau)$, as well as the new proposed $MTIE$ -based metrics (minimum peak-to-peak phase offset $mTIE(\tau)$ and the maximum frequency value $\hat{Y}_{pp}(\tau)$) was performed. In the presence of frequency synchronization, $MTIE(\tau)$ and $mTIE(\tau)$ indicate the relative value of the random phase deviation, neglecting the constant component x_0 . Thus, in the case of phase and ToD synchronization the current phase error $\Delta\varphi(t)$ is not less than $TIE(t, \tau)$, i.e. $\Delta\varphi(t) \geq x(t) = y_0 \cdot t$. In general, $\Delta\varphi(t)$ changes proportionally to $TE(t)$ and the normalized frequency value: $\Delta\varphi(t) \sim x(t) \sim y(t)$. So the following sequence of synchronization parameters evaluation is logical: FS \leftrightarrow PS \rightarrow ToD.

The Integration of Maximum Likelihood Estimation and the extended pick-to pick metrics (such as $MTIE/mTIE$) with PTP retains for further study.

The simulations and measurements performed on the operating networks has shown that to achieve the best result (carrier class TN quality), synchronization of modern packet networks should be carried out in the following sequence: (1) frequency synchronization of all network elements with the necessary accuracy at the given observation intervals τ ; (2) phase synchronization to align the phase offset between the

network elements at the given time intervals τ ; (3) time of day (ToD) synchronization at the given time intervals τ .

It should be considered that at the given observation intervals τ the relative accuracy of phase and time synchronization does not exceed the normalized accuracy of frequency synchronization.

Acknowledgements. Experimental studies and computer modeling were carried out with the support and useful communication of S. Mikhaylichenko and A. Korzh (Vodafone Ukraine), N. Legotin and N. Mokhovikov (ALTO Ltd, Russia) and Dr. Andreja Jarc (Meainberg, Germany). Also, the authors would like to thank M. Koltunov, M. Shwartz and A. Ryzhkov (MTUCI, Russia), D. Shevchenko (MTS Russia), D. Schneuwly and S. Kiss (Oscilloquartz, Switzerland) for their well-timed support and useful discussions.

References

1. Biriukov, N., Steklov, V.: Transport telecommunication networks: multiplexing systems. Student manual, 352 p. Vitol, Kyiv (2003)
2. Biriukov, N., Steklov, V., Kostic, B.: Transport telecommunication networks: multiplexing systems. Student manual for technical universities, 312 p. Technica, Kyiv (2005)
3. Bregni, S.: Synchronization of Digital Telecommunication Networks, p. 395. Wiley, Hoboken (2002)
4. Blechman, I.: Synchronization in science and technics, 544 p. Nauka, Moscow (1981)
5. Pikovsky A., Rosenblum, M., Kurths, Y.: Synchronization: A Universal Concept in Nonlinear Sciences, 496 p. Cambridge University Press, Cambridge (2002)/Technosphaera, Moscow (2003)
6. Biriukov, N., Triska, N.: Synchronization of telecommunication networks based on the synchronous and asynchronous transmission ways. *Electrosviaz* **10**, 34–37 (2007)
7. Biriukov, N., Triska, N.: Synchronization of transport technologies during the transition to NGN. *Electrosviaz* **9**, 30–35 (2009)
8. Triska, N.: Transport networks before and after NGN: the analysis of network synchronization aspects. *Zvyazok* **4**, 10–14 (2009)
9. Biriukov, N., Makurin, M., Triska, N.: The actual synchronization problems in the carrier-class networks. *Naukovi zapisky UNDIZ* **2**(10), 53–59 (2009)
10. Biriukov, N., Konovalov, G., Triska, N.: The actual trends of the development and standardization of frequency and time supporting in telecommunications. *Electrosviaz* **10**, 45–48 (2011)
11. Biriukov, N., Semenکو, A., Triska, N.: The problems of frequency and time supporting of the current telecommunications. *Computerni tehnologiyi drukarstva*, vol. 27, pp. 208–220. *Zbirnyk naukovykh pratz*, Ukrainian Academy of Printing, Lviv (2011)
12. Triska, N.: Synchronous ethernet as a synchronization method for packet networks. *Naukovi zapisky UNDIZ* **3**(11), 65–71 (2009)
13. Biriukov, N., Triska, N.: Synchronous ethernet as a basis for frequency and time supporting in the current and future telecommunications. *Electrosviaz* **2**, 8–12 (2013)
14. Biriukov, N., Triska, N., Khudyntsev, M.: The summary of ITU-T research activities in the field of frequency and time supporting of the modern telecommunication network. *T-Comm: Telecommun. Transport* **2**, 12–17 (2014)



15. Triska, N.: The actual issues of time synchronization during 5G networks deployment. In: 11th International Scientific-Technical Conference “Telecommunication problems”, 18–21 April 2017, ITS NTUU “KPI”. Conference Proceedings, pp. 102–104 (2017)
16. Triska, N.: Synchronization in new generation telecommunications: summary of technologies and standards. In: 12th International Scientific-Technical Conference “Telecommunication problems”, 16–20 April 2018, ITS, pp. 101–103. Igor Sikorsky Kyiv Polytechnic Institute. Conference Proceedings (2018)
17. Triska, N.: The guidelines for practical lessons on the academic subject “the fundamentals of synchronization networks and time distribution theory”: slips in the digital telecommunication systems and slip preventing methods. ITS NTUU “KPI”, Chair of Telecommunications, Kyiv (2015)
18. Biriukov, N., Triska, N.: Slip analysis in the plesiochronous network. *T-Comm: Telecommun. Transport* **10**(2), 19–24 (2016)
19. Biriukov, N., Triska, N.: The probabilistic analysis of information transmission in the plesiochronous network. In: 4-th International Scientific Conference “Infocommunications – Present and Future”, 30–31 October 2014, vol. 1, pp. 115–119, Odessa National O.S. Popov Academy of Telecommunications. Conference Proceedings (2014)
20. Biriukov, N., Triska, N.: The features of plesiochronous transmission in NGN and FN networks. In: 9th International Scientific-Technical Conference “Telecommunication problems”, 21–24 April 2015, pp. 125–127. ITS NTUU “KPI”. Conference Proceedings (2015)
21. Biriukov, N., Triska, N., Shwartz, M.: The packet delay estimation based on the packet transmission features. In: XII International Industrial Scientific-Technical Conference “Technologies of information society”, 14–15 March 2018. MTUCI. Conference Proceedings (2018)
22. Biriukov, N., Kilchychkiy, Y., Schmaliy, Y., Savchuk, A.: The problems of synchronization of the Ukrainian national telecommunication network. *Zbirnyk naukovykh pratz UNDIZ* **1**, 33–36 (1998)
23. Biriukov, N., Schmaliy, Y., Savchuk, A.: Telecommunication network synchronization in Ukraine. *Zvyazok* **3**, 47–50 (1998)
24. Biriukov, N., Khylenko, V., Kopijka, O., Suvorova, N.: Synchronization networks structure: the evolution of network synchronization approaches. *Zvyazok* **4**, 7–12 (2003)
25. Biriukov, N., Triska, N.: Synchronization networks: scenarios of interaction. *Seti i telecommunicatzii* **08–09**, 38–45 (2005)
26. Biriukov, N.: The design of potentially stable timing distribution networks. *Electrosviaz* **10**, 57–61 (2008)
27. Biriukov, N., Liskovskiy, I., Fedorova, N.: The method of analysis of synchronization networks sustainability. *Zvyazok* **3**, 23–27 (2006)
28. Biriukov, N., Liskovskiy, I., Fedorova, N.: The algorithm of loop detection in synchronization network. In: 7th International Scientific-Practical Conference “Modern information and electronic technologies”, 22–26 May 2006, Odessa, vol. 1, p. 190. Conference Proceedings (2006)
29. Biriukov, N., Triska, N.: Use of maximum time interval error function for estimation of digital paths condition. *Zvyazok* **1**, 28–32 (2004)
30. Biriukov, N.: Use of the TIE metrics for the evaluation of two-way precise time stamp protocols. In: *Zbirnyk naukovykh pratz VITI NTUU “KPI”*, vol. 1, pp. 24–29 (2011)

31. Biriukov, N., Triska, N.: The properties and applications of peak-to-peak time interval error metrics. In: Proceedings of Systems of Signal Synchronization, Generating and Processing in Telecommunications (SINKHROINFO 2017) Conference (IEEE Conference Record № 41975, Kazan), IEEE Digital Library: INSPEC Accession Number: 17080281. <https://doi.org/10.1109/sinkhroinfo.2017.7997502>
32. Biriukov, N., Triska, N., Khudyntsev, M.: Evaluation of the two-way time stamps exchange in packet network. In: 2017 4th International Scientific-Practical Conference “Problems of Infocommunications. Science and Technology” (PIC S&T-2017), 10–13 October 2017, Kharkiv, pp. 342–345 (IEEE Conference Record № 42018). Conference Proceedings (2017)
33. Biriukov, N., Makurin, M., Triska, N.: Analysis of the use of two-sided protocol for clock signals adjustment by time signals. *Naukovi zapisky UNDIZ* 4(16), 5–11 (2011)
34. Biriukov, N., Triska, N.: Analysis of slip frequency in a plesiochronous connection. In: 8-th International Scientific Conference “Infocommunications – Present and Future”, 14–16 November 2018, pp. 342–345. Odessa National O.S. Popov Academy of Telecommunications. Conference Proceedings (2018)

Radio Engineering and Electronics



Software Defined Radio in Communications

M. M. Kaidenko^(✉)  and D. V. Roskoshnyi^(✉) 

National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Peremoga Avenue 37, Kyiv 03056, Ukraine
{kkk610, ddd610}@ukr.net

Abstract. The abstract should summarize the contents of the paper in short terms, i.e. 150–250 words. The chapter covers Software defined radio systems and their usage in telecommunications. The presented results are based on relevant studies of authors conducted at the Institute of Telecommunication systems of “Igor Sikorsky Kyiv Polytechnic Institute”. Features of the modern Software-defined radio systems were analyzed. The architecture of Software defined radio system that is based on architecture of communication software was substantiated. The architecture that is oriented on realization of Cognitive Radio System was proposed. The creation process of Software defined radio system based on SDR Transceiver and technologies of System-on-chip using modern means of software oriented modeling was described in details. The variants of interaction between processor cores and programmable logic in the system on the crystal during creation of Software defined radio system were considered. The publication is addressed to scientists, graduate students, and developers of radiocommunication equipment.

Keywords: Software defined radio · Cognitive Radio System · SDR architecture · System-on-chip · Hardware-in-the-Loop Simulation

1 Introduction

1.1 Features of Modern Software Defined Radio

Software-Defined Radio (SDR) is a system that uses a set of hardware and software technologies, with some, or all functions of the radio system on a physical level implemented with the help of software. The software may be modified or may be embedded and operates on programmable signal processing devices. These devices can include field-programmable gate array (FPGAs), digital signal processors (DSPs), general purpose processors (GPPs), system-on-chip (SoC) systems, and other specialized programmable processors. The use of these technologies in building an SDR allows one to add new wireless features and capabilities to existing radio systems without changing the hardware platform.

Software-defined radio systems combine advanced technologies to provide flexibility in radio systems creation. These technologies include broadband technologies, radio frequency components, programmable radio frequency components, digital signal processing technologies, analog-to-digital (ADC) and digital-to-analog (DAC) converters,

frequency synthesizers, technology for building a system on a crystal, and information processing technology.

Due to technology advancements, SDR provides the opportunity to improve the use of the spectrum, opens up new telecommunications markets, and expands the list of available services, in particular, services related to new technologies for building the Cognitive Radio System.

Due to the flexibility of SDR technology, a number of problems arise in the configuration and management of the system: the assembly of equipment and software components of various platforms into a working radio system; control of the radio system during its operation; the work of the radio system in the conditions of physical and regulatory restrictions of the regulatory authorities of different countries; software self-configuration and adaptation of radio components; definition of common interfaces for applications and management software; the use of component libraries and radio configurations for various hardware and software platforms (portability).

1.2 Cognitive Radio Systems

Cognitive radio systems are perceived as a possible future solution due to the low availability of the radio frequency spectrum. It is a key technology that could provide reliable, flexible and efficient access to the spectrum by adapting mobile communication functions to the environment. The rapid and significant development of radio engineering in particular, Software Defined Radio (frequency and power management) contributed to the development of the cognitive radio. Cognitive radio can be characterized by the use of methods such as spectrum distribution in real time, broadband sensing, and real-time measurements. This revolutionary technology is a paradigm for changing the design of wireless systems since it will allow the rapid and efficient use of the radio frequency spectrum by offering distributed terminals the ability to communicate, self-regulate, and dynamically distribute the spectrum. The use of the spectrum can be greatly improved, allowing the secondary user to use the licensed band when the main user is absent. The cognitive radio, equipped with sounding and adaptation to the environment, is able to fill the holes in the spectrum and serve its users without creating harmful interference to the licensed user. For this purpose, the radio must constantly experience the spectrum it uses to detect the main user. After detecting the main user, the radio receiver must abandon the spectrum to minimize the obstacles that it may cause [1].

Some definitions of cognitive radio can be found in the research area. The official definition of cognitive radio systems in ITU, developed by ITURWP1B in 2009 and published in (ITU-R 2009), suggests that a cognitive radio system is a radio system that: uses technology that allows the system to gain knowledge of its operational geographic environment, the established policy, and its internal state; dynamically and autonomously adjust their operating parameters and protocols in accordance with the knowledge obtained in order to achieve predetermined goals; and learn from the results [2].

In other words, when cognitive radios can find opportunities to use “spectral holes” for communications, it is advisable to use cognitive radio communication to transport packets to facilitate the use of programs and services. A portable terminal with cognitive radio functions can collect information about the communication medium (for

example, spectrum holes, geographic location, access to wired/wireless communications or services available); analyze and study information from the environment with settings and user requirements; and to reconfigure by adjusting the system parameters to fit certain policies and norms. A cognitive terminal can also interact with other users of the spectrum and/or networks to provide a more efficient use of the network. The interaction procedure may be facilitated by network/infrastructure support or simply initiated in a special way [3].

2 The Architecture of a Software Defined Radio System for the Development of Radio-Telecommunication Equipment

The SDR software architecture was typically based on the Software Communications Architecture (SCA) [4], which was developed under the Department of Defense's Joint Tactical Radio Systems (JTRS) program [5]. SCA was called to provide an open architecture that allows radio engine developers to combine hardware and software components into one system. In essence, SCA might be perceived as an "operating system" that defines the interaction of SDR software components. However, SCA transitions to SDR from software point of view, and not from the implementation of the radio system and the prospects for its life cycle point of view. Although the SCA approach provides flexibility for software components, the transition from software components to working radios is not unambiguous. Solving this ambiguity is very important in the creation of congruence radio systems, in [6] authors proposed the architecture model of a software defined radio system for the development of equipment.

The architecture presented on Fig. 1 will solve the problems of ambiguity for developers of modern radio systems, in particular, the cognitive radio systems.

The proposed architecture is based on the SCA specification and was used to identify the required components for the real cognitive radio system described [7, 8]. One of the main components used in this model is the library manager that was used to provide organized administration of the components for the many files that are required for the SCA program. The model also demonstrates the use of component classification, including the corresponding Application Programming Interface (API). The API describes the components for the project in a way that allows the components to have an interface with a common architecture. Such approach improves the system's self-awareness, which in turn improves cognitive capabilities of the system, allowing the radio system to determine which services provide an accessible components and how to access them simply by knowing the type of classification.

The architecture uses the Intelligent Controller to communicate with the applications level. It also includes a manager that defines a set of rules for defined control policies. Policies and rules are generated for a specific radio device assignment within the components available for its implementation. The intelligent controller determines which features are available and which components are active or that need to be

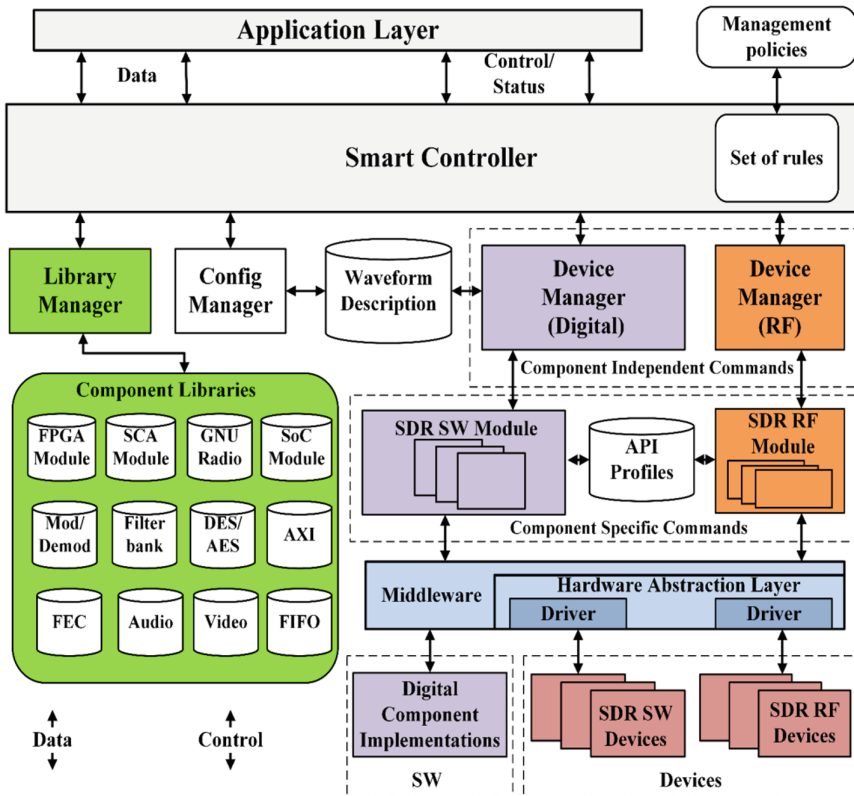


Fig. 1. Architecture model of a program-determined radio system for the development of radio-telecommunication equipment of the cognitive radio systems.

activated. Additionally, the Intelligent Controller must include an application-level access control module to provide protection against unauthorized access.

The architecture provides two separate device managers for managing digital devices and radio frequency devices. This allows independent control of devices and improves portability from platform to platform. Device managers create an operational radio system, the information on which is described in the software application description (SAD). Dynamic radio systems can contain several SADs that allow the system to be compatible with different standards and different transmission profiles for a single standard. After SAD analysis, digital devices and radio frequency devices are configured with the device manager in accordance with the description. Information for setting up devices (profiles) is stored in the component library. This library stores registered components that are available to the radio system and is managed by the library manager, which is responsible for adding components to the library, removing them and connecting them. The components library is used not only to support the work of the radio system and to improve the organizational structure of radio software, but also to improve its autonomy. In the case of a cognitive radio controlled by the

Intelligent Controller, the most optimal option for constructing the radio system in accordance with the task is determined. When creating a real radio system that is to adapt as a cognitive radio system, or to use adaptation to change the conditions of work in the presence of non-stationary noise [9], the intelligent controller, in accordance with the generated rules, transfers their set to the controller (processor) of the system control together with the specified components.

Logical devices control their hardware devices through the appropriate interface. By using the API classification, the logical device should only know the type of device, while giving access to the corresponding basic function. This allows a logical device, such as an amplifier, to control it as long as it is used by the device driver. For a particular component, a logical device is a simple template that will not be needed to create a system in real time.

3 The Process of Creating a Software-Defined Radio System

Creating and prototyping wireless systems have been discussed for many decades, but only in recent years the possibility of a complete development cycle on a PC - from the creation of the model to the completion of the implementation emerged. This is due to the development of Matlab/Simulink software-oriented simulation tools and their integration with FPGA design, digital signal processors, controllers, and embedded systems using high-level programming languages.

With the advent of the system on the crystal technology, even more opportunities were created for the development of SDR through model-oriented design with the use of high-level software design. Usage of SoC, which combines the versatility of the ARM processor and the power of FPGA processing, allows to allocate resources for implementing high-speed and low-speed digital signal processing processes, decision-making algorithms and control algorithms between processor cores and FPGA. High-speed digital processing tasks that require real-time execution, such as modulation, jamming coding, synchronization, channel status estimation, and others, are performed on FPGA. For data decoding, visualization, monitoring and diagnostics, support for external hardware and software interfaces processor cores are used.

General process of creating a radio system using model-oriented design is shown in Fig. 2.

The first step is to create a model of the radio system in the Matlab/Simulink environment, while the model must have modules that are responsible for interacting with the hardware platform (SDR model). When creating a model, one takes into account what part of the model (algorithms) will be implemented on the FPGA or on the ARM processors. The part of the model that will be implemented on the ARM processor is software, and it includes all the architectural modules of the program defined radio system that will be implemented programmatically.

The debugging of the model in the first stage is carried out in the mode of programmed hardware simulation. Hardware-in-the-Loop Simulation (HIL) is a method used in the development and testing of complex technical and technological embedded real-time systems. The use of software and hardware modeling allows one to conduct in-circuit and intra-system testing of developed solutions and mathematical models and

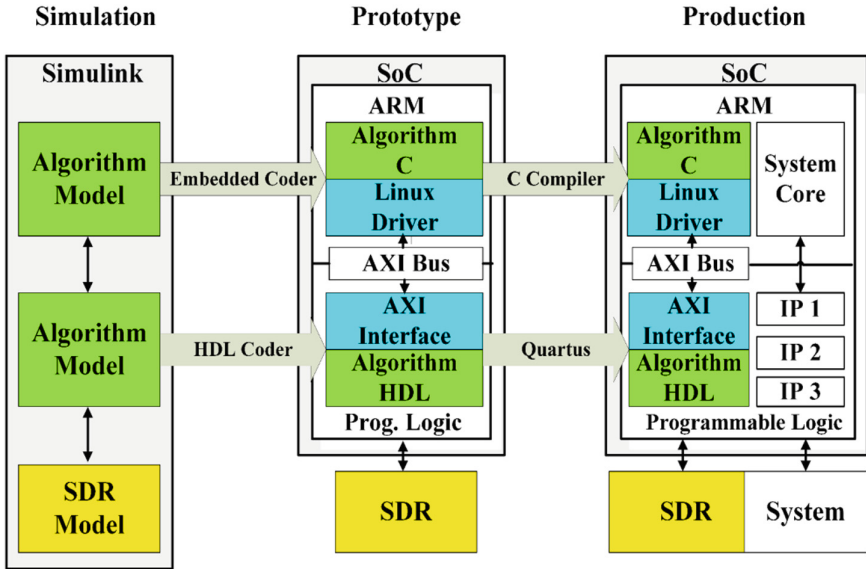


Fig. 2. The structure of the process of creating a program-defined radio system using model-oriented design.

significantly reduce the timing and cost of development. The process of HIL - simulation in general consists of three main stages:

- creating a mathematical model of the real environment in which the hardware device will be used;
- testing of the developed mathematical model on a real hardware device;
- implementation of the process in a real device, taking into account the results of testing the model.

Usage of HIL simulations open the possibility to use different options using a mathematical description (software model) of the environment and in-circuit simulation of the process being developed, as well as the mathematical description of the process (software model) and in-system simulation of its behavior in the real environment.

Modern SDR and SoC technologies that were included in the Matlab/Simulink support packages for HIL simulation of telecommunication processes and systems was significantly expanded. One of the latest solutions is to use the SDR technology Analog Devices based on AD9361 in conjunction with SoC Altera or Xilinx.

Software-hardware modeling of telecommunication processes and systems in Matlab is based on the use of the IIO System Object [10], which in turn is based on the specification of Matlab's system objects. The IIO System Object works both with Matlab and Simulink and is intended for data exchange on the Ethernet network with an ADI hardware platform, connected to the FPGA/SoC platform running on the ADI Linux operating system. The IIO System Object is built on the libiio library [11] and

allows the Matlab or Simulink model to exchange data flows with the hardware platform, monitor its settings, and monitor various parameters.

To use the Matlab environment for work with the software-hardware platform, a configuration file with .cfg extension is created with a name that matches the name of the device installed in the block configuration dialog. This file contains a set of fields that defines Linux drivers for the target device and configuration channels that will be used by the Simulink block or Matlab script.

On the second stage, after the model has been worked out, HDL code (IP Core) is generated in VHDL (Verilog) using the HDL Coder and C code using the Embedded Coder, which are a part of the Matlab/Simulink package. At this stage, debugging is done using the hardware platform for prototyping the SDR, which consists of the SDR module of the radio transceiver and the SoC module is carried out. The debugging is carried out in Simulink “External” simulation mode. The effectiveness of the developed algorithms and models is checked, all the necessary settings are implemented, the system is already working both under the control of Matlab/Simulink and autonomously.

The final step is the automatic compilation of the code using the gcc linux compiler and VHDL compiler (Verilog), which is part of the Quartus package (for SoC Altera).

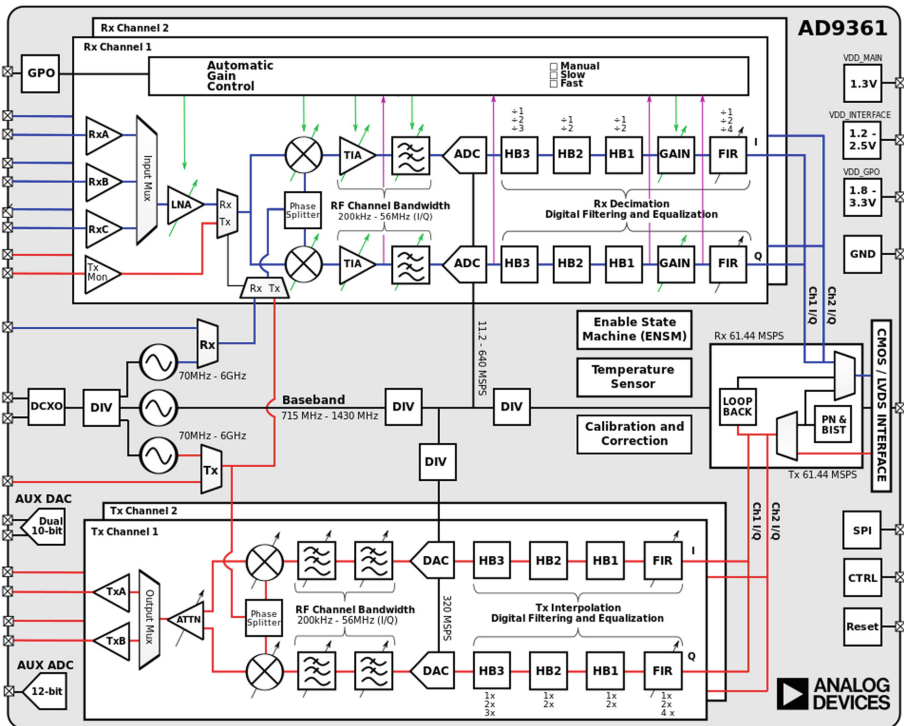


Fig. 3. Structure of the SDR transceiver AD9361

The process of creating a radio system with the use of model-oriented design ensures that once the algorithm of the SDR system is implemented in the final product, it will be fully tested and ensures the reliability of the system.

The process of creating the radio system has been tested [7]. Testing was performed using the Altera SoC Cyclone V and HSMC ARRADIO SDR Module based on AD9361 from Analog Devices [12], the structure of which is shown in Fig. 3. SDR transceiver AD9361 is designed for use in base stations of 3G and 4G networks. Transceiver's receiver and transmitter are constructed according to the scheme of direct frequency conversion. Using direct conversion of the frequency requires a high degree of filtration on the receiver side channels and transmitter out-of-band radiation. To do this, the transceiver employs the required number of digital and analog filters, which parameters are installed programmatically. The transceiver also includes frequency synthesizers, a reference oscillator, low noise amplifiers, power amplifiers, an automatic gain control system, and a software management interface. The transceiver is fully program-driven, which allows you to create SDR systems of various purposes, including the radio-based radio systems.

4 Interaction Between HPS and FPGA SoC During Creation of Radio Telecommunication Device

The development and usage of systems-on-crystal (SoC) is one of the most promising areas in the development of telecommunication devices and systems. In general, SoC is a further development of FPGA technology and is a hardware processor core (or cores) integrated in a single crystal, a FPGA programmable logic matrix, and a hardware or software implemented peripheral device management modules. The internal data bus with high bandwidth of more than 10 Gbit/s, which connects the processor core with the FPGA logic matrix, allows you to manage peripheral modules and in some cases accelerate the execution of algorithms, thus increasing the overall system performance. This level of integration not only increases productivity, but also reduces power consumption, size and cost of the final solution [13].

Communication applications, including wireless systems and its SDR solutions, use HPS + FPGA technology to accelerate the development of end-to-end telecommunication devices, the output of which can be obtained by a powerful device capable of providing most of the required computing in real-time (due to the use of FPGA), which can be easily integrated into any system due to the ability to use such interfaces as Ethernet, I2C, SPI, CAN and others, at a higher level of the operating system.

The main feature of SoC is the availability of programmable blocks. Therefore, SoC is not just an integrated circuit, but a compound, which includes both a hardware part (actual FPGA and peripheral devices), and a software program that includes embedded software. Thus, when designing SoC, it is necessary to perform operations for joint verification and debugging of the interaction between the software and hardware components. An additional advantage of using SoC technology in SDR systems is that it allows designing a data link at the structural level, while low-level designing will only be needed to implement the most cost-efficient solution in terms of the computing resource [14].

Two main approaches are used while developing SoC systems:

- creation of a device without an operating system (Bare Metal Application);
- usage of the operating system VxWorks/Linux.

The first approach allows you to get a system with the maximum level of performance, but its implementation is extremely complicated, as it involves the full development of all necessary devices, interfaces, etc. for SDR at a low level.

The second approach facilitates the task, since the operating system is managing peripherals, implementing high-level exchange protocols such as IP, TCP, UDP and even device access protocols such as FTP, NTP, Telnet, SSH, etc. Also, using the operating system allows you to create universal devices that are much easier to integrate into existing systems than when using Bare Metal Application.

The most expedient way to create new systems is to use the Linux operating system for SoC devices, since it has the following advantages [15]:

- extensive hardware and software support for the developer;
- huge flexibility in applications on the finished system;
- easy to use peripherals;
- availability of broad support of development tools for embedded systems;
- support for multi-core processors and ARM processors at the kernel level.

There are three basic approaches on how to organize HPS and FPGA interactions in Linux:

1. Use direct access to FPGA memory registers.
2. Use the user interface for I/O.
3. Create a full-fledged Linux device driver on a functional level.

The first approach is the easiest way to exchange data between the FPGA and the processor bypassing the operating system, which nonetheless makes it possible to use all of its other capabilities. Interaction with FPGA occurs through direct write/read to the memory addresses that the operating system provides. It does not allow you to use the advanced features of the operating system and the FPGA itself, such as Direct Memory Access (DMA).

In the second approach, the I/O interface of the Linux operating system allows you to create a simplified version of the drivers for sharing data between the FPGA and the CPU. This approach provides the ability to work with interruptions in the operating system, but there is no ability to control caching, which prevents using the technology of direct access to memory.

The third approach is comprehensive, and it allows you to use all the capabilities of the operating system and FPGA when working with SoC with any peripherals, including ADC and DAC in SDR. This approach also allows usage of one driver for devices of the same type, which greatly simplifies the process of development of the finished systems.

Regardless of the HPS and FPGA interaction approach used, the use of SoC devices in conjunction with SDRs can significantly accelerate the development of wireless telecommunication devices and systems.

At testing and approbation of the process of creating a radio system using SDR and SoC, a full-fledged approach was used. Figure 4 shows the structure of the interaction organization between HPS and FPGA SoC when working with SDR transceiver AD9361. The operating system provides support for standard kernel drivers and user created drivers which provide interaction between user applications and IP Core AD9361 to ensure the management of the SDR transceiver through the AXI Lite bus.

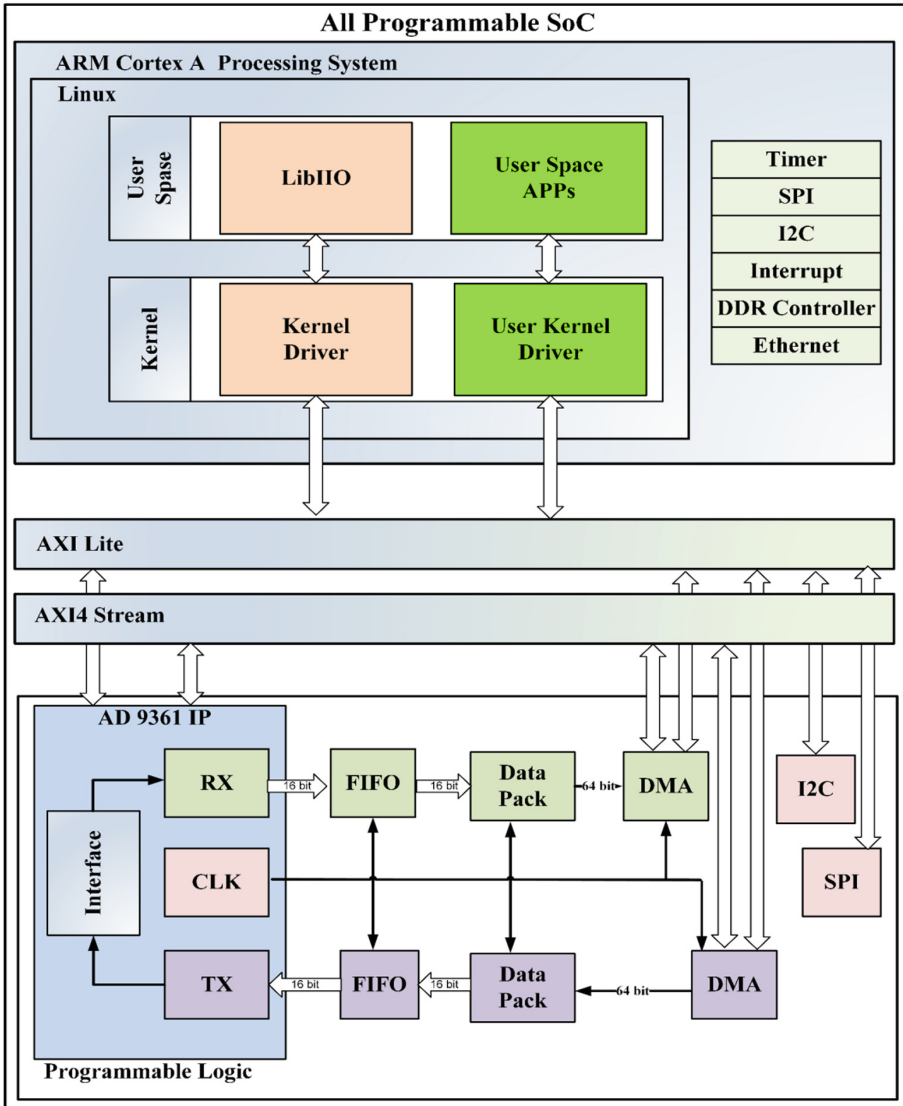


Fig. 4. Structure of interaction between HPS and FPGA SoC when working with SDR transceiver AD9361.

High-speed ADC data transfer, AD9361 DAC and FPGA data transmitted via the AXI 4 bus, use standard kernel drivers and the libiio library. The data exchange between HPS and FPGA is done via DMA in 64-bit format, which is converted into packets of 16-bit data for exchange through FIFO buffers with IP Core AD9361.

5 Conclusion

This paper presents the results of the work of authors on the creation of software defined radio systems for telecommunications. The developed architecture of a software-defined system is based on Software Communications Architecture Specification. It focuses primarily on SDR system developers and takes into account the peculiarities of such systems from the process of creating the system and the prospects for further development of the system and its accompaniment in the process of life cycle points of view. The proposed solution on transfer of generated rule set in the controller (processor) which controls the real SDR system that should adapt as a congenial system, or adapt to changing conditions of work in conditions of non-stationary noise [9], is original and allows us to effectively use the software resource and at the same time to carry out the modernization of the system during the life cycle without changing the hardware platform.

The presented process of modern SDR system creation is based on the experience of the authors' usage of several modern technologies. SoC technology, which allows you to combine versatility of the ARM processor and the power of FPGA processing, compromisingly allocating resources for implementation of high-speed and low-speed processes of digital signal processing, decision-making algorithms and algorithms between processor cores and FPGA. Technology of modern single-chip SDR transceivers creation allows you to create fully programmable control SDR system, which is very important when creating a system of radio-sensing. The technology of model-oriented design using high-level design software tools, combined into a single set of tools aimed at creating the final product with minimal time and cost. The creation process was tested when creating a real SDR system [7].

An important aspect when using SoC technology is a compromise decision on the organization of interaction between the processor cores and FPGA. The results of the analysis are presented by the authors of various approaches and a validated solution on the structure of the organization of interaction between the HPS and the FPGA SoC with a robot with an AD9361 SDR transceiver is presented.



Further work will be aimed at creating a real cognitive SDR system with multi-level adaptation to changing conditions in the presence of unintended and deliberate interference caused by various interfering factors.

References

1. Letaief, K.B., Zhang, W.: Cooperative communications for cognitive radio networks. *Proc. IEEE* **97**(5), 878–893 (2009)
2. Towards Cognitive Radio Systems. Main Findings from the COGNAC project. In: *Matinmikko, M., Bräysy, T. (eds.) Espoo 2011. VTT Tiedotteita - Research Notes 2575*
3. Chen, K.C., Peng, Y.J., Prasad, N., Liang, Y.C., Sun, S.: Cognitive radio network architecture: part I - general structure. In: *Proceedings of the 2nd International Conference on Ubiquitous Information Management and Communication (ICUIMC 2008), Malaysia (2008)*
4. Software Communications Architecture Specification. Final, 15 May 2006. Version 2.2.2. http://www.public.navy.mil/jtnc/SCA/SCAv2_2_2/SCA_version_2_2_2.pdf. Accessed 14 Jan 2019
5. Joint Tactical Radio Systems. <http://jtrs.army.mil>. Accessed 10 Mar 2017
6. Newman, T., Minden, G.J.: A software defined radio architecture model to develop radio modem component classifications. In: *First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, DySPAN 2005*, pp. 1–4. IEEE Xplore Digital Library (2005). <https://doi.org/10.1109/dyspan.2005.1542674>
7. Ilchenko, M., Kaidenko, M., Kravchuk, S., Khytrovskyy, V.: Compact troposcatter station for transhorizon communication. In: *Proceedings of the 2017 International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo) 11–15 September 2017*, pp. 1–4. IEEE Conference Publications, Odessa, IEEE Xplore Digital Library Ukraine (2017). <https://doi.org/10.1109/ukrmico.2017.8095364>
8. Kravchuk, S., Kaidenko, M.: Features of creation of modem equipment for the new generation compact troposcatter stations. In: *Proceedings of the International Scientific Conference “RadioElectronics & InfoCommunications” (UkrMiCo 2016), 11–16 September 2016*, pp. 365–368. IEEE Digital Library, Kyiv, Ukraine (2016)
9. Kaydenko, N.: Adaptive modulation and coding in a broadband wireless access systems. In: *23th International Crimean Conference on “Microwave and Telecommunication Technology”, CriMico 2013, Sevastopol, 8–13 September 2013*, pp. 275–276 (2013)
10. ADI vs MathWorks Support. https://wiki.analog.com/resources/tools-software/linuxsoftware/libiio/clients/fmcomms2_3_simulink. Accessed 14 Jan 2019
11. What is libiio? <https://wiki.analog.com/resources/tools-software/linux-software/libiio>. Accessed 14 Jan 2019
12. AD9361 RF Agile Transceiver. <http://www.analog.com/en/products/rf-microwave/integrated-transceivers-transmitters-receivers/wideband-transceivers-ic/ad9361.html#product-overview>. Accessed 14 Jan 2019
13. Accelerate Internet of Things (IoT) Innovation. https://www.altera.com/content/dam/altera-www/global/en_US/pdfs/literature/br/br-industrial-brochure.pdf. Accessed 14 Jan 2019
14. Intel SoC FPGA Embedded Development Suite User Guide. https://www.intel.com/content/dam/www/programmable/us/en/pdfs/literature/ug/ug_soc_edu.pdf. Accessed 14 Jan 2019
15. Customizable ARM Designs and Linux. <https://www.denx.de/en/pub/Documents/Presentations/2013-06-25-customizable-arm.pdf>. Accessed 14 Jan 2019



Application of Spatial Signal Processing by the Form of the Electromagnetic Wave Phase Front in Wireless Communication Systems

Hlib Avdieienko^(✉)  and Yevhenii Yakornov^(✉) 

National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic
Institute”, Peremoga Avenue 37, Kyiv 03056, Ukraine
djangov2006@ukr.net, yakornovits@gmail.com

Abstract. The main purpose of this chapter is to familiarize readers with such new trend of wireless communication systems development and improvement as spatial signal processing by the form of the electromagnetic wave phase front. The issues of spatial signal processing practical application are represented by investigations for improving functional capabilities of the phase direction-finders and phase systems for coordinates determination, first of all, the possibility of distance determining to radio-frequency source which locates in the Fresnel region. The features of spatial signal processing application by the form of the electromagnetic wave phase front for the future development of radio relay communication lines which ensure reuse of frequency band are shown. The prototype description of the simplex radio relay communication line is presented. This prototype provides frequency band reuse factor of 2, i.e. transmission and reception of two digital radio signals at the data rate of 41 Mbps in same frequency band of 8 MHz due to the spatial signal processing application by the form of the electromagnetic wave phase front.

Keywords: Spatial signal processing · Direction-finder · Antenna array · Phase front · Electromagnetic wave

1 Introduction

In modern radio engineering systems (e.g., air traffic control systems, multiposition radar and phase direction-finding systems, satellite communication systems, etc.) spatial signal processing is an integral part of the space-time signal processing.

Space-time signal processing (STSP) is a set of operations with radio signals received in different points of space by antenna elements or antenna subsystems located at one or more stations. STSP is performing to extract desired information in the form of a message or measurement of one or several signal parameters in order to determine spatial location of the radio-frequency source (RS), a speed and direction of their movement, etc. The main content of the STSP is a synthesis of optimal in one way or another signal processing systems and the analysis of the quality of these systems [1, 2].

In most cases, the space-time narrowband signal condition is fulfilled for wireless communication systems (WCS), so STSP can be divided into two independent stages. The first is a spatial signal processing stage and the second is time processing stage. This, in turn, allows to ensure ensuring the possibility of WCS technical construction for STSP implementation.

In the presence of external radio frequency interferences (RFI) which coincide over the frequency band and electromagnetic wave (EMW) polarization radiation with the frequency band and the EMW radiation polarization of a desired signal, a weight summation of the output voltages from the antenna system (AS) elements is carried out at the stage of spatial signal processing. Such spatial signal processing provides an optimal compromise between spatial accumulation of the desired signal components and the mutual compensation of external RFI components. So, a compromise between amplification of the AS in the angular direction of the desired signal RS and its weakening (“nulling”) in the angular directions of the interference RS is done. Moreover, the greater the intensity of the external RFI in comparison with the intensity of a receiver internal noise, the deeper nulling in the AS radiation pattern in the angular directions of interference RSs. Optimal time processing is synthesized regardless of spatial signal processing, that is, assuming that the configuration of the AS is completely set.

The application of adaptive STSP provides high efficiency of interferences suppression when spatial locations of interference RSs differ from spatial location of the desired signal RS.

The main disadvantage of STSP is the inability or low effectiveness of interference suppression coming from angular directions close to the direction of arrival of the desired signal RS, especially when they fall into the radiation pattern main lobe of the AS. In this case, spatial signal processing by angular coordinates is ineffective and for the desired signal separation against the RFI it is necessary to use the difference in other features of the desired signal and the interference. The form of the electromagnetic wave (EMW) phase front of the desired signal and RFI on the aperture of the receiving AS may be such feature.

In general, we can distinguish three main areas of spatial signal processing application in WCS by the form of the EMW phase front [3–7]:

- (1) spatial separation of the desired signal against the interference background in case of angular coordinates coincidence of the desired signal RS and the interference RS which located in different wave radiation regions;
- (2) position finding of the desired signal RS or interference RS (the bearing angle and the distance) located in the Fresnel region by application of phase direction-finding system;
- (3) ensuring the frequency resource reuse for the radio relay communication line (RRCL) by forming several independent spatial channels of the EMW transmission in the same radio frequency band and with the same EMW polarization but with the different form of the EMW phase front. This ensures data throughput increase of this frequency band for RRCL.

The first direction of spatial signal processing application by the form of the EMW phase front theoretically is well described in a number of references, for example in [1, 2]. Therefore, the scientific novelty of the research is to consider ways of spatial signal processing application by the form of the EMW phase front for the phase-direction finding systems and RRCL.

2 General Information About the EMW Phase Front

As a rule, there are three main types of the EMW phase front: spherical, cylindrical and plane. The spherical phase front usually has EMW emitted from the point RS and cylindrical phase front has EMW emitted from linear RS. The EMW which are emitting by any antenna and propagating in a homogeneous isotropic medium has spherical phase front (spherical waves) at a large distance from it. Thus, the EMW near the Earth's surface which radiated by the communications satellite antenna can be considered as spherical waves. Also EMW received by the antenna of the radio relay station (RRS) No. 1 emitted by the antenna of the RRS No. 2 which communicate each other are spherical waves. But this applicable only for the condition that the distance between these RRS is much greater than the antenna linear size. In both cases, the transmitting antenna can be considered as point emitter of the EMW with spherical phase front, since its linear size are much smaller than the length of the distance between the radio transmitter and the radio receiver device.

For case of aperture antennas (parabolic, horn, antenna arrays, etc.) application which generate directed EMW radiation in the radio links, depending on the form of the EMW phase front of RS, it is customary to distinguish between three radio wave radiation regions: the near-field (NF) region, the intermediate or Fresnel region, far-field (FF) or Fraunhofer diffraction region (see Fig. 1) [1].

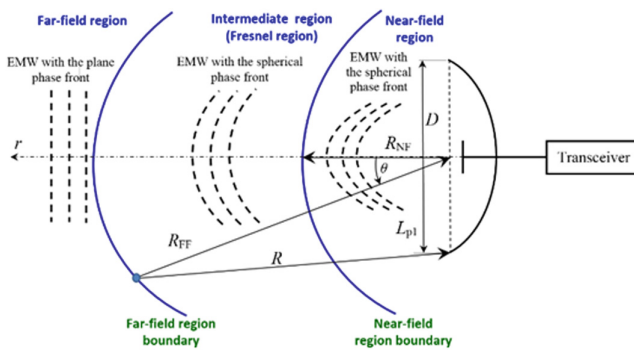


Fig. 1. Graphic explanation for the delimitation of radio wave radiation regions

It is assumed [1] that a signal from the RS located in the FF region forms an EMW field distribution with the plane or quazi-plane phase front at the receiving antenna aperture. When the RS of the signal is located in the NF region or in the Fresnel region it forms EMW field distribution with the spherical phase front at the receiving antenna aperture.

In turn, the boundaries of the NF and the FF regions are determining by the next relations [1, 3, 9, 10]

$$R_{NF} = 0.6\sqrt{\frac{D^3}{\lambda}}, \quad R_{FF} = \frac{2\pi D^2 \cos^2 \theta}{\lambda} \quad (1)$$

where:

λ - the wavelength,

D - the antenna aperture linear size (Fig. 1),

θ - the angular direction of EMW reception (or the EMW radiation in case of the transmitting antenna), counted from the axis or phase center of the antenna.

Thus, if the desired signal RS and the interference RS under the condition of their angular coordinates coincidence are located in different wave radiation regions of the receiving antenna (for example, the desired signal RS located in the FF region and the interference RS located in the NF region or vice versa) then desired signal at the antenna aperture will forms EMW with plane phase front and the interference will forms EMW with spherical phase front or close to it. This difference in the forms of the EMW phase front can be used in the receiving device to separate desired signal and suppress the interference, or vice versa to separate the interference and suppress the desired signal by using spatial signal processing by the form of the EMW phase front.

3 Spatial Signal Processing by the Form of the EMW Phase Front Application in the Phase Direction-Finding Systems

The main purpose of the phase direction-finder application is to determine the bearing angle of RS location. Most of variants of the phase direction-finders developed for the present time determine only the RS bearing angle for the EMW with the plane form of the phase front, that is, for case when the RS located in the FF region of the phase direction-finder antenna array (AA). The simplest phase direction-finder for the determining RS bearing angle β by the EMW with the plane phase front usually contains a linear sparse 2-element AA (exact, but ambiguous scale L_{43} for the phase shift measurement), in the center of which there is a 2-element linear AA (rough, but unambiguous phase shift measurement scale L_{21}), as shown in Fig. 2.

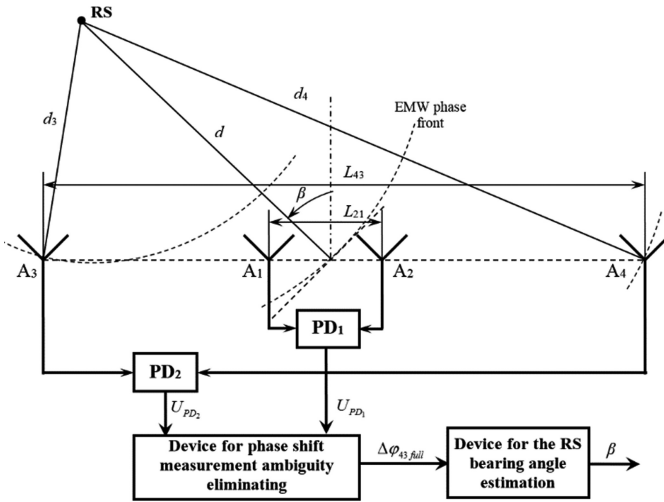


Fig. 2. The block diagram of the simplest phase direction-finder: PD is a phase detector

In the paper [4] it was proven that the phase direction-finder shown in Fig. 2 also allows determining the RS bearing angle, which radiate EMW with the arbitrary form of the phase front, including the spherical phase front.

However, the phase direction-finder application in the super-high frequency range can lead to a situation when the RS located into Fresnel region of the phase direction-finder AA. That is RS stay from AA at the distance R that satisfies the condition

$$R_{NF} < R < R_{FF}, \tag{2}$$

The phase direction-finder, shown in Fig. 2, does not allow determining the distance to the RS which is located in the Fresnel region. Therefore, in order to solve this problem, it is necessary to increase the number of antenna elements of sparse AA of phase direction-finder up to 3. Under the condition (2) all known phase direction-finding systems which use three-element sparse AA and operate under the assumption of presence of the EMW with the plane phase front will inevitably make significant errors in determining the bearing angle of RS which radiates EMW with the spherical phase front.

At the same time, theoretical possibility of determining the distance to RS and the RS bearing angle by the form of the EMW spherical phase front in the phase direction-finding systems which use the 3-element sparse AA is shown in the paper [3]. But firstly, the mechanism for the eliminating phase shift measurement ambiguity between the adjacent antenna elements of the AA is ignored. Secondly, there is lack of mechanism for determining the type of EMW phase front, which necessary for the direction-finder system switching to the other mode of mathematical calculation of the RS bearing angle compared with the mode of RS location in FF region.

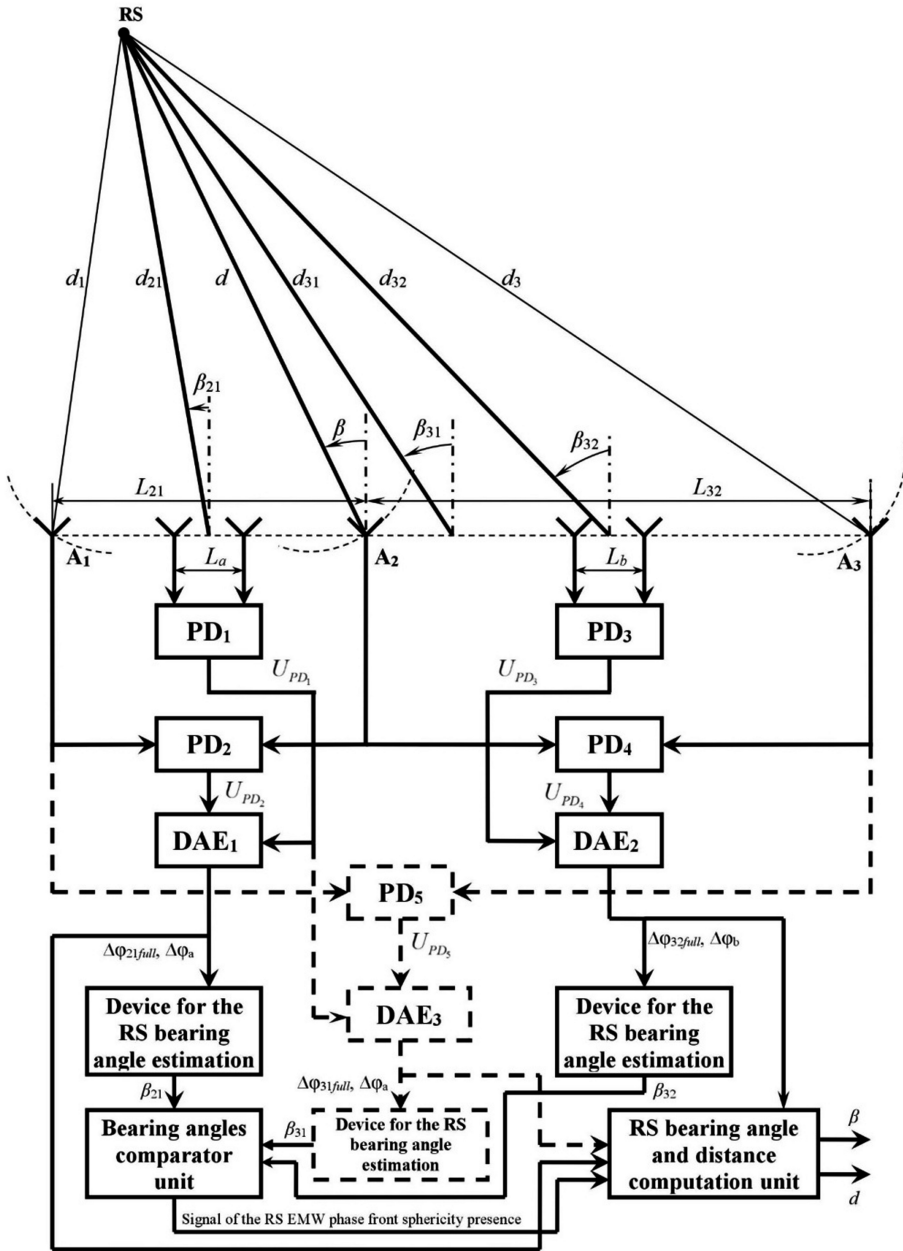


Fig. 3. The phase direction-finder based on three element sparse AA and two unambiguous phase shift measurement scales: DAE is a device for ambiguity eliminating, PD is a phase detector

Therefore, consider the technical solution for the phase direction-finder constructing with the application of the 3-element sparse AA, which allows determining the RS bearing angle in case of its location in the FF region. Also this technical solution allows determining the RS bearing angle and the distance to RS in case of its location in the Fresnel region of the direction-finder AA (see Fig. 3). This direction-finder implies the use of two precise but ambiguous phase shift measurement scales L_{21} and L_{32} respectively, as well as two rough but unambiguous phase shift measurement scales $L_a = L_b = 0.5\lambda$ which located within corresponding ambiguous scales [4].

In case when the EMW with plane phase front from RS incidents on the AA of such phase direction-finder, then according to Fig. 3 the following relationships are true

$$\beta = \beta_{21} = \beta_{32} = \beta_{31}, \quad (3)$$

$$\Delta\varphi_{21full} = \Delta\varphi_{32full} \frac{L_{21}}{L_{32}} = \Delta\varphi_{31full} \frac{L_{21}}{L_{31}}, \quad (4)$$

$$\begin{aligned} \Delta\varphi_{21full} &= \Delta\varphi_{21meas} + 2\pi k_{21}, \Delta\varphi_{32full} = \Delta\varphi_{32meas} + 2\pi k_{32}, \Delta\varphi_{31full} \\ &= \Delta\varphi_{31meas} + 2\pi k_{31}, \end{aligned} \quad (5)$$

$$\begin{aligned} k_{21} &= \left\langle \frac{1}{2\pi} \left(\frac{L_{21}}{L_a} \Delta\varphi_a - \Delta\varphi_{21meas} \right) \right\rangle, k_{32} = \left\langle \frac{1}{2\pi} \left(\frac{L_{32}}{L_b} \Delta\varphi_b - \Delta\varphi_{32meas} \right) \right\rangle, \\ k_{31} &= \left\langle \frac{1}{2\pi} \left(\frac{L_{31}}{L_a} \Delta\varphi_a - \Delta\varphi_{31meas} \right) \right\rangle, \end{aligned} \quad (6)$$

where:

$\Delta\varphi_{21full}$, $\Delta\varphi_{32full}$, $\Delta\varphi_{31full}$ - complete differences of the EMW phase, respectively, between the 2nd and the 1st, the 3rd and the 2nd, the 3rd and the 1st antennas of the direction-finder AA,

$\Delta\varphi_{21meas}$, $\Delta\varphi_{32meas}$, $\Delta\varphi_{31meas}$ - differences of the EMW phase, respectively, between the 2nd and the 1st, the 3rd and the 2nd, the 3rd and the 1st antennas of the direction-finder AA. They are measured with usage of the phase detectors (PD) PD₂, PD₄, PD₅,

k_{21} , k_{32} , k_{31} - coefficients which taking into account the lost number of complete signal periods during phase difference measurements for the corresponding phase direction-finder ambiguous scales,

$\langle \cdot \rangle$ - rounding operation to the nearest integer value,

$\Delta\varphi_a$, $\Delta\varphi_b$ - phase shifts which are measured for corresponding unambiguous scales L_a and L_b .

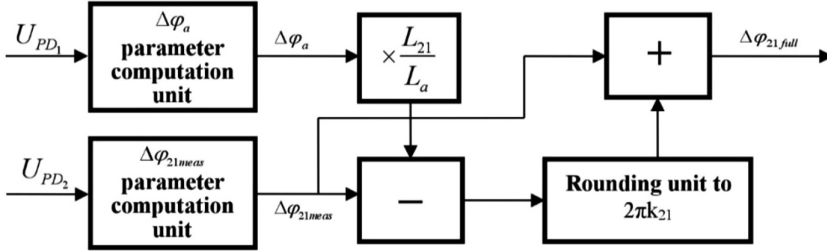


Fig. 4. The block diagram of the device for ambiguity eliminating of phase shift measurement $\Delta\varphi_{21full}$

In Fig. 3 the dashed line shows the bearing angle β_{31} measurement scheme elements for the measurement scale $L_{31} = L_{21} + L_{32}$, which may not be used if the information obtained by phase difference measurements for the scales L_{21} and L_{32} will be sufficient. A simplified, typical block diagram of the device for ambiguity eliminating of phase shift measurement $\Delta\varphi_{21full}$ is shown in Fig. 4.

Thus, if one of the Eqs. (3) or (4) is not satisfied then we can assume that the EMW phase front from the RS at the aperture of the phase direction-finder AA is spherical front. Then, for both the plane and the spherical front of the EMW, we can write that [4]

$$\begin{aligned} \beta_{21} &= \arcsin\left(\frac{\lambda}{2\pi L_{21}} \Delta\varphi_{21full}\right), \beta_{32} = \arcsin\left(\frac{\lambda}{2\pi L_{32}} \Delta\varphi_{32full}\right), \\ \beta_{31} &= \arcsin\left(\frac{\lambda}{2\pi L_{31}} \Delta\varphi_{31full}\right). \end{aligned} \tag{7}$$

In the presence of the EMW with spherical phase front at the aperture of the phase direction-finder AA, under the assumption that the RS located in the Fresnel region and the distance between the RS and the AA location is greater than the AA linear size of the direction-finder, that is $L_{31}/d < 1$, the following approximations are true

$$d_1 \approx d - L_{12} \sin \beta + \frac{L_{12}^2 \cos^2 \beta}{2d}, d_3 \approx d + L_{32} \sin \beta + \frac{L_{32}^2 \cos^2 \beta}{2d}. \tag{8}$$

Then, the complete phase differences $\Delta\varphi_{21full}$, $\Delta\varphi_{32full}$ can be expressed in the form

$$\begin{aligned} \Delta\varphi_{21full} &= \frac{2\pi(d - d_1)}{\lambda} = \frac{2\pi}{\lambda} \left(L_{12} \sin \beta - \frac{L_{12}^2 \cos^2 \beta}{2d} \right), \Delta\varphi_{32full} = \frac{2\pi(d_3 - d)}{\lambda} \\ &= \frac{2\pi}{\lambda} \left(L_{32} \sin \beta + \frac{L_{32}^2 \cos^2 \beta}{2d} \right), \end{aligned} \tag{9}$$

Solving together the Eq. (9), we obtain that the desired bearing angle β and the distance d to the RS due to the sphericity of the EMW phase front, will be equal to

$$\beta = \arcsin \left[\frac{\lambda L_{21} (\Delta\varphi_{21full} + \Delta\varphi_{32full})}{2\pi(L_{21} + L_{32})L_{32}} \right], \quad (10)$$

$$d = \frac{L_{21}^2 \left[4\pi^2 (L_{21}^2 + L_{32}^2) L_{32}^2 - \lambda^2 L_{21}^2 (\Delta\varphi_{21full} + \Delta\varphi_{32full})^2 \right]}{4\pi\lambda(L_{21} + L_{32})L_{32} \left[L_{21}^2 \Delta\varphi_{32full} + \Delta\varphi_{21full} (L_{21}^2 - L_{32}^2 - L_{21}L_{32}) \right]}. \quad (11)$$

The main significant disadvantage of the phase direction-finder block diagram which shown in Fig. 3 is the necessity of the application of the two unambiguous scales for bearing angle measurement. It is equivalent to the additional use of four near-omnidirectional antennas. In addition, due to the ultimate overall dimensions of the near-omnidirectional antennas in the microwave frequencies band which exceeding 0.5λ , there may be a situation where it will be impossible to place them side by side at a distance of $L = 0.5\lambda$.

The solution to this situation may be application of the phase direction-finder scheme with three-element AA, where the formation of unambiguous scales for bearing angle measurement is carried out through the frequency division circuits usage (Fig. 5).

It is worth noting that the frequency division schemes usage are possible only in case of the narrowband or harmonic signal radiation from the RS. In this case, for the formation of unambiguous but coarse measurement scales of the phase shifts in the phase detectors PD₁ and PD₃, the frequency division coefficients must satisfy equations

$$n_{21} = \frac{2L_{21}}{\lambda}, \quad n_{32} = \frac{2L_{32}}{\lambda}. \quad (12)$$

The main disadvantage of the above-mentioned phase direction-finder construction (Fig. 5) is significant complication due to frequency division schemes application.

Thus, for direction-finding of the RS located in the Fresnel region by the EMW phase front curvature by applying the phase direction-finder consisting of three-element sparse AA, two operations must be performed sequently: (1) to determine the presence of the EMW phase front sphericity; (2) to switch the phase direction-finder from the bearing angle evaluation mode for the EMW with plane phase front into the mode of bearing angle assessment for the EMW with spherical phase front. Also the obligatory phase shifts measurement ambiguity elimination must be completed as shown above.

If can add another linear AA which will be orthogonal to the phase direction-finder AA shown in Fig. 3 or Fig. 5, the overall AA becomes planar AA. It makes possible to measure not only the RS bearing angle β but also the RS tilt angle α . Such phase direction-finder is called the phase system for position finding (PSPF). Let us consider, in accordance with [5], the simplified scheme of the plane AA arrangement (Fig. 6), which can be supplemented by unambiguous measurement scales b_1, b_2, b_3 to eliminate the ambiguity of the phase shifts measurements.

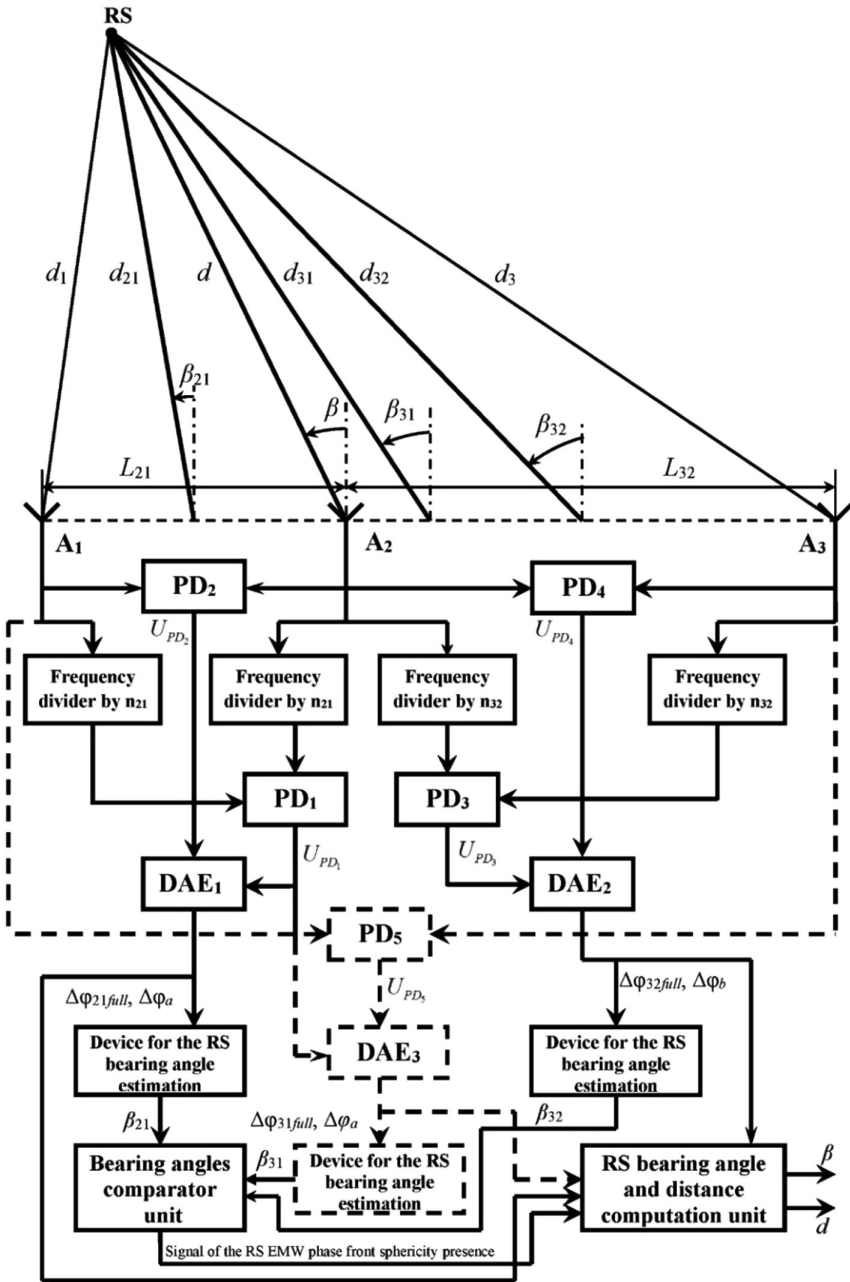


Fig. 5. The phase direction-finder which uses the EMW phase front curvature of the RS located in the Fresnel region. It based on 3-element sparse AA and frequency division circuits joint application

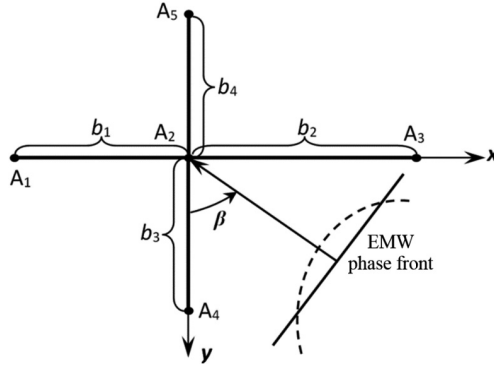


Fig. 6. Antennas arrangement scheme in the planar AA of the PSPF

In case where the EMW phase front of the form the RS at the PSPF aperture of AA has plane form, the bearing (β) and tilt (α) angle of the RS site are calculated by using the next formulae

$$\beta = \arctg \left[\frac{2b_3 \arctg(F(\beta, \alpha))}{(b_1 - b_2) \Delta\varphi_{42full}} \right], \alpha = \arccos \sqrt{\left(\frac{\lambda \arctg(F(\beta, \alpha))}{\pi(b_1 - b_2)} \right)^2 + \left(\frac{\lambda \Delta\varphi_{42full}}{2\pi b_3} \right)^2}, \quad (13)$$

where:

$F(\beta, \alpha) = \operatorname{tg} \left[\frac{\pi(b_1 - b_2)}{\lambda} \cos \alpha \sin \beta \right]$ - the direction-finding characteristic for the EMW with plane phase front,

$\Delta\varphi_{42full} = \Delta\varphi_{42meas} + 2\pi k_{42}$ - complete phase shift between the antennas A_4 and A_2 of the planar AA,

$\Delta\varphi_{42meas}$ - measured phase shift between the antennas A_4 and A_2 of the planar AA, k_{42} - the lost integer number of signal periods during phase difference measurement by the phase detector.

In the presence of the EMW phase front sphericity, the angular coordinates of the RS are calculated according to the next formulas

$$\beta = \arctg \left[\frac{(b_3 b_4^2 + b_3^2 b_4)(b_2^2 \Delta\varphi_{21full} + b_1^2 \Delta\varphi_{32full})}{(b_1 b_2^2 + b_1^2 b_2)(b_3^2 \Delta\varphi_{52full} + b_4^2 \Delta\varphi_{42full})} \right], \quad (14)$$

$$\alpha = \arccos \sqrt{\left(\frac{\lambda(b_2^2 \Delta\varphi_{21full} + b_1^2 \Delta\varphi_{32full})}{2\pi(b_1 b_2^2 + b_1^2 b_2)} \right)^2 + \left(\frac{\lambda(b_3^2 \Delta\varphi_{52full} + b_4^2 \Delta\varphi_{42full})}{2\pi(b_3 b_4^2 + b_3^2 b_4)} \right)^2}, \quad (15)$$

where:

$\Delta\varphi_{52full} = \Delta\varphi_{52meas} + 2\pi k_{52}$ - complete phase shift between the antennas A_5 and A_2 of the planar AA,

$\Delta\varphi_{52meas}$ - measured phase shift between the antennas A_5 and A_2 of the planar AA,
 k_{52} - the lost integer number of signal periods during phase difference measurement by the phase detector.

The distance to the RS in the presence of the EMW phase front sphericity can be calculated using the following approximate formula

$$d \approx \frac{\pi b_4^2 (1 - \cos^2 \alpha \sin^2 \beta)}{\lambda \Delta\varphi_{42full} - 2\pi b_4 \cos \alpha \sin \beta}. \quad (16)$$

So, in determining the angular coordinates of the RS located in the Fresnel region for the EMW with the spherical phase front on the basis of the planar AA of PSPF, it is necessary to perform two sequential operations: (1) to determine the presence of the EMW phase front sphericity; (2) to switch PSPF from the bearing and the tilt angles of the RS evaluation for the EMW with the plane phase front to the mode of the bearing and tilt angles of the RS assessment, as well as, if necessary also distance to the RS for the EMW with the spherical phase front.

4 Application of Spatial Signal Processing in the Radio Relay Communication Lines

Nowadays, the main trend in the WCS development is the expansion of number and quality of the various information and communication services (telephony, access to the Internet, transmission of multimedia information in real time, etc.) provision to users. However, the deployment of such services is realized mainly due to the technical ability to increase the data transmission speed by subscriber channel within frequency band, which allocated for a specific wireless communication system. Therefore, the task of ensuring frequency resources rational usage by telecommunication operators due to radio-frequency spectrum limitation and competition for it between different radio technologies does not lose its importance to this day.

The mentioned task is related to the RRCLs, which are well-proven in the wireless segments construction of different operators' telecommunications networks at backbones, intra zonal and local levels, as well as in the deployment of radio links between mobile cellular systems base stations and so on.

The cardinal solution of the radio-frequency spectrum deficiency is the conversion of millimeter (30–300 GHz) and terahertz bands. For example, commercial proposition of the RRCL equipment of E-band (71–76 and 81–86 GHz) has emerged on the telecommunication market and provides maximum data transmission speed up to

5 Gbps in 500 MHz radio channel with distance between radio-relay station of 5 km [6]. Also, there are experimental prototypes of RRCL equipment of terahertz band (130–134 GHz) which showed the ability to transmit data with speeds of several Gbps [7, 8]. However, it should be noted that cost of such equipment is still very high compared to cost of the RRCL in the microwave band.

All aforementioned points out the reasonability for the further searching of the ways of radio-frequency spectrum rational use for RRCLs by improving the existing radio signals separation methods (spatial, polarization, frequency, time, code separation, etc.). By applying this method in practice on the receiving side of RRCL, it will be possible to separate different subscribers radio signals with use the same frequency band and EMW polarization from each other with minimal distortion of these signals and their mutual influence on each other. At the same time, it is highly desirable that in practice, these methods will require minimal expenditures to replace the existing hardware and software solutions of existing and future RRCLs.

In the authors opinion of this research, one of the promising ways that can be used to solve the problem of frequency band reuse for RRCL, which in turn will increase RRCL bandwidth capacity, is spatial signal processing by the form of the EMW phase front.

Taking into account the urgency of the solution of the frequency bands lack problem for RRCLs which operate below 40 GHz, the theoretical study of the possibility of spatial signal processing application by form of the EMW phase front was given in the works [9–11] to ensure the reuse of the RRCL radio-frequency spectrum for its bandwidth capacity increase.

The main idea that explains spatial signal processing application by the form of the EMW phase front for RRCLs is presented in works [9, 10]. It consists in the following: by means of the beamforming circuits (BFC) of the transmission path of the radio relay station (RRS) No. 1 the radio signals from various data sources artificially transform to the EMWs. These EMWs radiate by the AA of RRS No. 1 in the angular direction of the RRS No. 2 in the same frequency band and the same polarization, but with different forms of the EMWs phase fronts in the free space between corresponding RRS No. 1 and RRS No. 2. Next, there is also subsequent spatial separation of the aforementioned EMWs from each other by forms of their EMW phase fronts in the AA and the BFC of reception path of the RRS No. 2.

Consider the generalized structural scheme (Fig. 7), which explains the idea of spatial signal processing application by the form of the EMW phase front for unidirectional (i.e. simplex) RRCL [12, 13].

Assume that the EMW transmission goes from RRS No. 1 in the direction of RRS No. 2. The signal from the i -th ($i = 1 \dots M$) transmitter of the RRS No. 1 enters at the input of the power divider of the i -th BFC, where occurs its division into M of identical signals. Next, these M signals from outputs of the power divider enter to inputs of corresponding channels of the i -th BFC, where they are multiplied by the complex weight coefficients. The multiplication changes the signals amplitude and the phase.

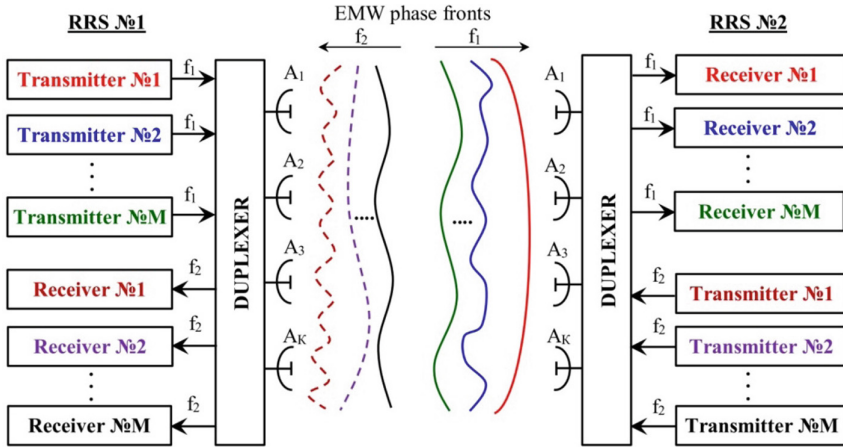


Fig. 7. The block diagram of one-hop RRLC with spatial signal processing application

From the i -th BFC outputs (Fig. 8a), M signals are fed into corresponding inputs of M combiners, where they combine with signals from other transmitters. Then the total signal comes to the bandpass filter, and then to the output of the l -th ($l = 1 \dots K$) AA antenna element of the RRS No. 1 (Fig. 9). As a result, in the transmission path with the help of the corresponding BFC the formation of EMW for M different transmitters is carried out. The EMWs emit by AA of RRS No. 1 in the same frequency band and with same type of polarization.

In this case, the distinctive feature of these radio signals are various forms of its EMW phase front at the AA aperture of the reception path. For example, EMW from the first transmitter of the RRS No. 1 has the plane phase front, the EMW from the 2nd transmitter of the RRS No. 1 has spherical or close to the spherical form of phase front and the EMW from the M transmitter has the phase front of the sinusoidal or parabolic form at the AA aperture of the RRS No. 2. At the receiving side spatial separation of the signals of different transmitters from each other by the EMW phase front form is carried out within corresponding BFC (Fig. 8b) by changing the signals amplitudes, phases and followed by its combining.

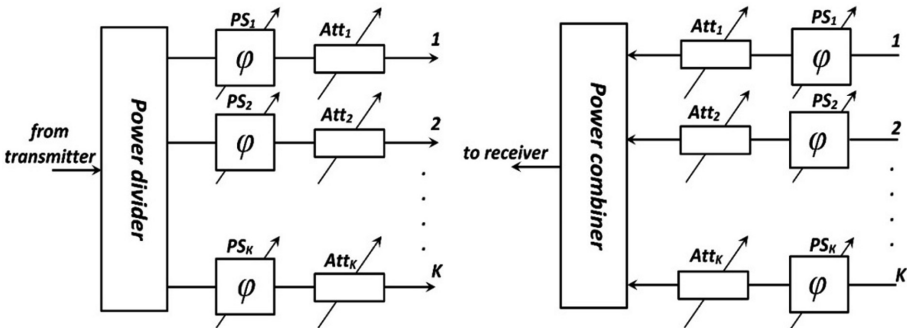


Fig. 8. The block diagrams of the BFC for the transmission (a) and the reception (b) path

As a result, it is clear that the key elements of spatial signal processing are the BFCs, which are EMW formers with different phase fronts in the transmission path of the RRS No. 1 (No. 2) and the EMW spatial filters by the form of the EMW phase front in the reception path of the RRS No. 2 (No. 1). The main characteristic of the BFC is a set of complex weight coefficients that reflect the settings of its phase shifters and attenuators (in case of the analog implementation of the BFC).

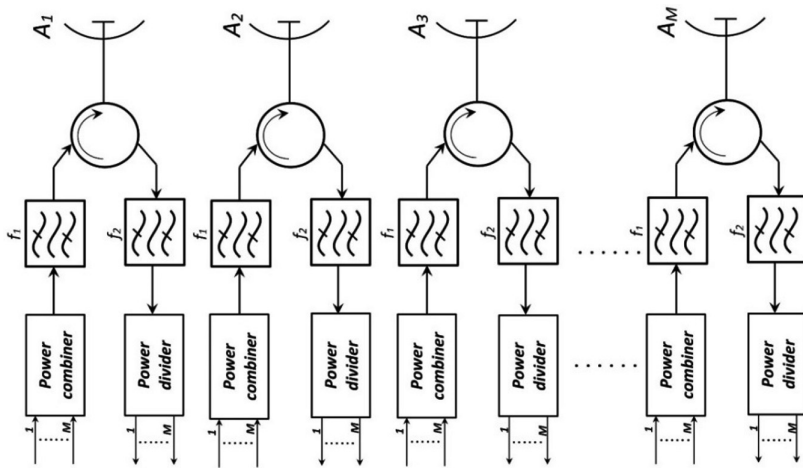


Fig. 9. The block diagram of the duplexer for the RRS with spatial signal processing in transmission and reception paths

Accordingly, the set of BFC weight coefficients for the transmission and reception paths should be such that to ensure the best signals separation of different transmitters of RRS No. 1 (No. 2) from each other in the reception path of the RRS No. 2 (No. 1). In this case, the effectiveness of spatial selection refers to the ability of the individual i -th BFC of the RRS No. 2 (No. 1) reception path to separate with the smallest attenuation signal with desired form of the EMW phase front and at the same time suppress the signals of transmitters from the RRS No. 1 (No. 2) with the other forms of the EMW phase fronts.

5 A Prototype of the Simplex RRCL with Spatial Signal Processing

Block diagrams of transmitting and receiving stations of the proposed prototype of the simplex RRCL are presented in Figs. 10 and 11.

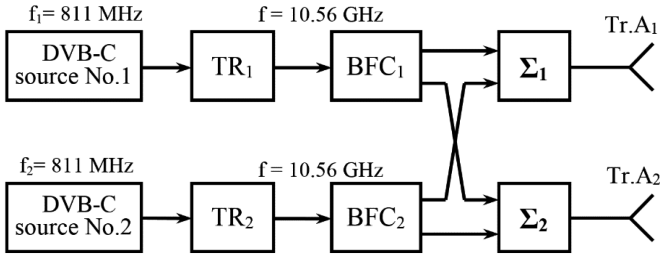


Fig. 10. The block diagram of the transmitting station of the proposed simplex RRCL

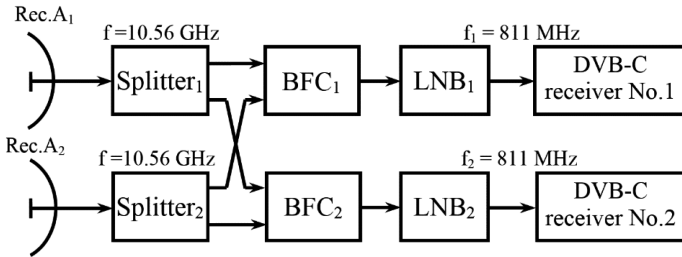


Fig. 11. The block diagram of the receiving station of the proposed simplex RRCL

In compliance with Fig. 10 transmitting station of proposed RRCL consists of two DVB-C signal sources, microwave transmitters (TR), BFC, combiners and horn antennas. The first DVB-C signal source is WISI OV75 transmodulator, which is tuned to receive DVB-S signal from Astra 4A satellite on geostationary earth orbit with $f = 11.766$ GHz carrier frequency, 36 MHz signal bandwidth, $V = 27.5$ Msps (i.e. 55 Mbps data rate) symbol rate and converts it to DVB-C signal with 64-QAM modulation, $V = 6.875$ Msps (i.e. 41.25 Mbps data rate) symbol rate, 8 MHz signal bandwidth and $f_1 = 811$ MHz output carrier frequency. Second DVB-C signal source consist of IRD-2600 CODICO Scopus satellite receiver and Radyne Comstream QAM-256 modulator. IRD-2600 receiver is tuned to receive DVB-S signal from Astra 4A satellite with $f = 12.073$ GHz carrier frequency, 36 MHz signal bandwidth, $V = 27.5$ Msps symbol rate. Radyne Comstream QAM-256 modulator converts MPEG transport stream from output of IRD-2600 modulator to DVB-C signal with 64-QAM, $V = 6.875$ Msps symbol rate, 8 MHz signal bandwidth and $f_2 = 811$ MHz output carrier frequency. The microwave transmitter consists of the frequency converter from 811 MHz to 10.56 GHz, bandpass filter and the two-stage transistor amplifier. The output power of the TR is 0.06 mW. The BFC of the transmitting station is assigned for control radiated EMW front curvature at the antenna aperture of receiving station by changing its amplitude and phase. The BFC is series connection of the waveguide power divider, the variable attenuator and the variable phase shifter (Fig. 12).

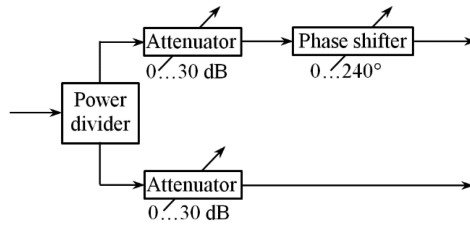


Fig. 12. The block diagram of the beamforming circuit of the transmitting station

Microwave signals from BFC is transmitted to input of combiners Σ_1 , Σ_2 and from its output to horn antennas of $15 \times 10 \text{ cm}^2$ aperture and with 21 dBi gain. Horn antennas simultaneously radiate electromagnetic waves of two DVB-C signals with the vertical polarization and 8 MHz frequency bandwidth toward receiving station of the RRCL. Figure 13 is the photo of the simplex RRCL transmitting station element.



Fig. 13. Photo of the transmitting station of the proposed simplex RRCL

In compliance with Fig. 11 the receiving station of the proposed simplex RRCL consists of two offset parabolic dish antennas, splitters, spatial filters, low-noise blocks (LNB) of 10.7–12.75 GHz band and two DVB-C receivers, which connected to television set. Dish antennas has diameter of 35 cm, which corresponds to 33 dBi gain at 11 GHz. This antennas are located along a straight line next to each other. The splitter is waveguide power divider with two outputs. Each output of the splitter is connected to the appropriate spatial filter.

The spatial filter is the series connection of the waveguide variable attenuator and the variable phase shifter which connected to two inputs of the waveguide power combiner (Fig. 14).

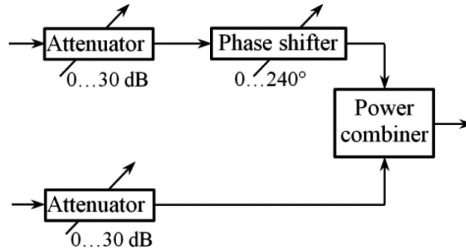


Fig. 14. The block diagram of spatial filter of the receiving station

The unit of the Spatial filter₁ is assigned to separate the first DVB-C source radio signal from the second DVB-C source radio signal with same frequency band and the polarization by using differences of the form of the EMW phase fronts from this sources. Similarly, Spatial filter₂ is assigned to separate second DVB-C source radio signal from first DVB-C source radio signal only using differences of the form of the EMW phase fronts of this sources. After spatial selection in spatial filters each radio signal comes to the LNB where it converts from 10.56 GHz to intermediate frequency 811 MHz. From LNBs output DVB-C radio signals comes to Homecast and Kaon DVB-C receivers where they demodulate and decode to TV programs. The photo of the simplex RRCL receiving station is shown in Fig. 15.

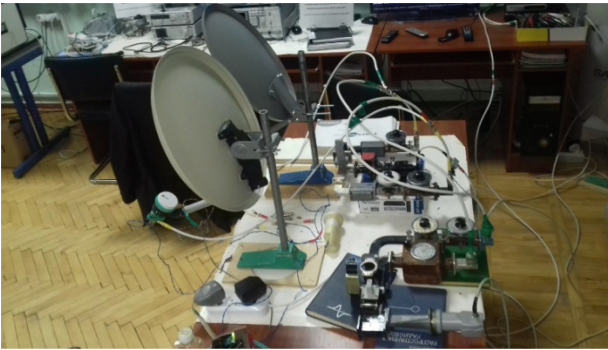


Fig. 15. The photo of the receiving station of the proposed simplex RRCL

The photo of spatial filters of the receiving station is shown in Fig. 16.

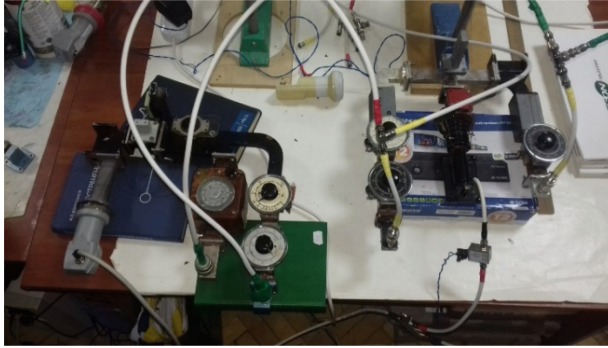


Fig. 16. The photo of spatial filters of the receiving device

6 Results of Simplex RRCL Prototype Lab Test

Lab test of the simplex RRCL prototype with spatial signal processing is accomplished. The distance between the transmitting and the receiving part is 7 m. Two EMW (which corresponds to the first and the second DVB-C signal) with different curvature of the phase front was realized by manual control of signal phases and its amplitudes in BFCs of transmitting station. In turn, spatial processing at the receiving station was realized by the manual control of variable attenuators and phase shifters of spatial filters. In order to separate first DVB-C source radio signal, attenuators and phase shifters of the Spatial filter₁ is manually adjusted by operator to maximum suppress power of second DVB-C source radio signal. Similarly, in order to separate second DVB-C source radio signal, attenuators and phase shifters of the Spatial filter₂ is manually adjusted by the operator to maximum suppress power of the first DVB-C source radio signal. Power spectrums of first and second DVB-C signals at the input of the DVB-C receiver No. 1 and No. 2 after the signal processing in the Spatial filter₁ and Spatial filter₂ respectively are presented as photos in Figs. 17 and 18.

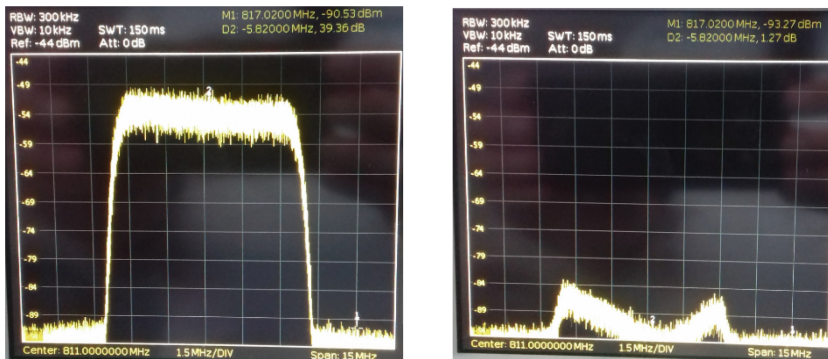


Fig. 17. The power spectrum of the first (a) and the second (b) DVB-C radio signal at the DVB-C receiver No. 1 input after spatial signal processing

From comparison of power spectrums of DVB-C signals at Fig. 17a it is obvious that power level of the first DVB-C signal is $P_1 = -51.2$ dBm, its out-of-band noise level $N_1 = -90.5$ dBm and its in-band noise level, which caused by second DVB-C signal residual after spatial signal processing is equals to $\Delta P_2 = -92...-84$ dBm (Fig. 17b) in 8 MHz bandwidth. So, the carrier-to-noise ratio of the first DVB-C signal is equals to $(C/N)_1 = 40.8$ dB at the carrier frequency of DVB-C and 32.8 dB at the edge of DVB-C signal spectrum. Better suppression of the second DVB-C signal in Spatial filter₁ was obtained at the edge of DVB-C signal spectrum.

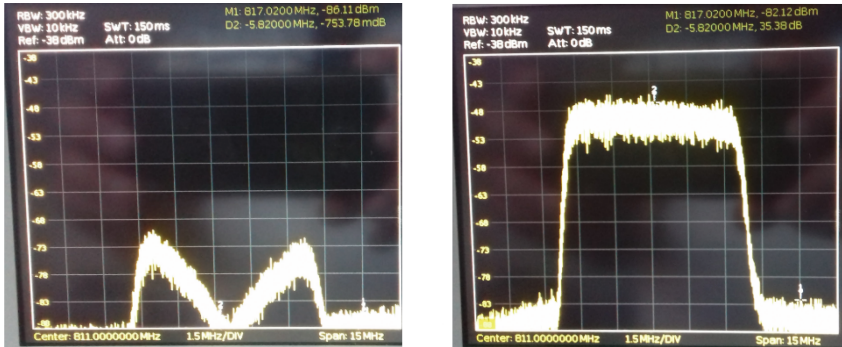


Fig. 18. Power spectrum of first (a) and second (b) DVB-C radio signals at the DVB-C receiver No. 2 input after spatial signal processing

Similarly, from comparison of power spectrums of DVB-C signals in Fig. 18b it is obvious that the power level of the second DVB-C signal is $P_2 = -46.7$ dBm, its out-of-band noise level $N_1 = -86.1$ dBm and its in-band noise level, which caused by the first DVB-C signal residual after spatial signal processing is equals to $\Delta P_1 = -88...-70$ dBm (Fig. 18a). So, the carrier-to-noise ratio of the second DVB-C signal is equals to $(C/N)_2 = 41.3$ dB at carrier frequency and 23.3 dB at the edge of DVB-C signal spectrum. Better suppression of the first DVB-C signal in the Spatial filter₂ was obtained at the center frequency of DVB-C signal spectrum.

The nonidentity of the residual spectrums of DVB-C signals shown in Figs. 17b and 18a after spatial signal processing at the receiving station of the RRCL prototype can be explained by the nonidentity of the technical characteristics of the devices of BFCs and spatial filters (i.e. variable attenuators, variable phase shifters, power dividers) in the frequency band of the DVB-C signal which is equals to 8 MHz.

The conducted experiment shows that the maximum rejection ratio of the interfering signal DVB-C in the corresponding channel of the receiving side does not exceed 40–45 dB at the center frequency and 28–35 dB at the edge of the signal spectrum. This is due to an inaccurate setting of the variable phase shifters and variable attenuators when they are mechanically tuned by the operator.

Constellation diagram for the first DVB-C signal before and after spatial suppression of the second DVB-C signal is presented in Fig. 19. Constellation diagram for the second DVB-C signal before and after spatial suppression of the first DVB-C signal

is presented in Fig. 20. These constellation diagrams were measured with application of CrazyScan software and the DVB-T/T2/C QBox TBS5880 receiver which was connected by turn to the LNB₁ and the LNB₂ of the receiving station of the RRCL prototype.

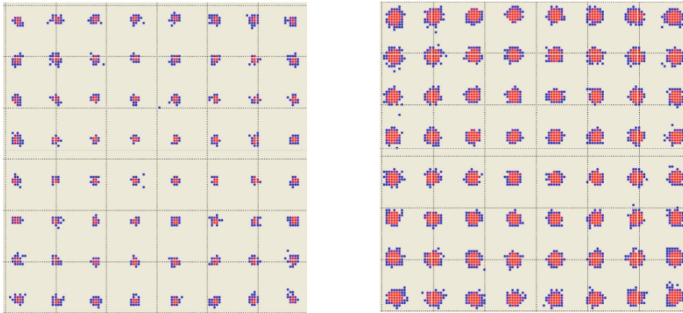


Fig. 19. Constellation diagram for the first DVB-C signal before (a) and after spatial suppression (b) of the second DVB-C signal

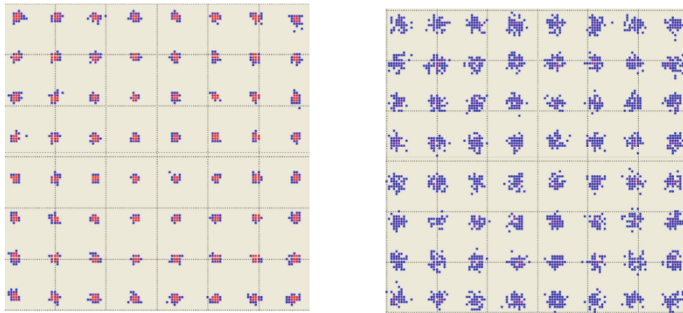


Fig. 20. Constellation diagram for the second DVB-C signal before (a) and after spatial suppression (b) of the first DVB-C signal

In accordance with the constellation diagram parameters measurement the signal-to-noise ratio (SNR) degradation and bit-error rate (BER) for the first DVB-C signal after spatial suppression of the second DVB-C signal is 3 dB and $\text{BER} < 10^{-8}$. SNR degradation and BER for the second DVB-C signal after spatial suppression of the first DVB-C signal is 8 dB and $\text{BER} = 3.4 \times 10^{-5}$.

From DVB-C receivers outputs composite video signal and two-channel audio connected to television set. DVB-C radio signals quality of demodulation and decoding in DVB-C receiver No. 1 (Kaon) and No. 2 (Homecast) is illustrated at Fig. 21.

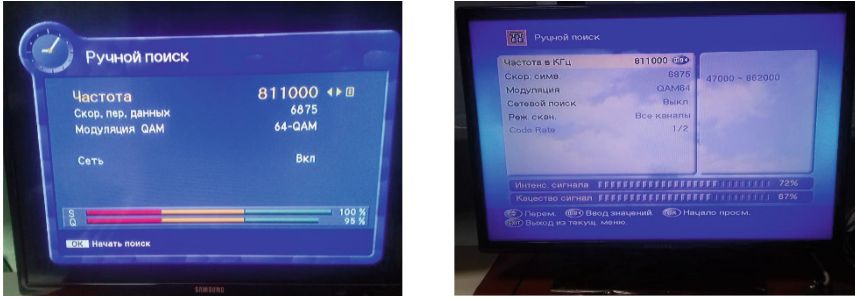


Fig. 21. The first (a) and the second (b) DVB-C radio signal levels (S), quality of demodulation and decoding (Q) in DVB-C receivers Kaon and Homecast after spatial signal processing

From Fig. 21 it is evidently that the quality of demodulated and decoded DVB-C radio signal in the corresponding DVB-C receiver after spatial signal processing not less than 95% for the first DVB-C signal and 67% for the second DVB-C signal. These levels of quality is sufficient in order to ensure high quality of video and audio services of the each television program without any interruptions at the TV set screen.

7 Conclusions

So, we can conclude, that during RS angle coordinates measurements based on the EMW phase front curvature with phase-direction finder application for case when RS located within the Fresnel region of phase-direction finder sparse AA it is sufficient to perform two operations:

- (1) determine whether the EMW phase front from RS is plane or non-plane;
- (2) switch the phase-direction finder operation mode for the appropriate form of the EMW phase front and, if needed, calculate distance from phase direction-finder sparse AA to RS based on the EMW phase front sphericity.

Knowledge of RS distance from AA may be used in passive radars, radio interference suppression systems for case when interference RS have the same bearing angle with desired RS, but located in different radiation regions (e.g. desired signal RS in the FF region and interference RS in the NF or Fresnel region or vice versa), in mobile communications systems to improve signal level in some areas, in dispatch control systems based on satellite navigation systems, etc.

The radio relay communication line with spatial signal processing in the transmitting and receiving devices based on differences in the form of EMW phase fronts is proposed. The prototype of simplex RRCL with spatial processing was developed and tested. Distance between the transmitting and the receiving device in lab test was 7 m. This prototype can simultaneously transmit and receive two DVB-C radio signals with 41 Mbps speed in the same frequency band and the polarization type. Demodulation and decoding quality of DVB-C radio signals after spatial signal processing in the RRCL prototype not less than 67%. This level of quality is sufficient in order to ensure

high quality of video and audio services of the each television program without any interruptions. So, the application of spatial signal processing can ensure radio frequency spectrum reuse factor of 2 for the RRCL.



Future direction of the research may be the study of spatial signal processing efficiency based on the form of the EMW phase front at longer distances between the transmitting and receiving devices of the RRLC and with number of DVB-C signals more than 2.

References

1. Kremer, I.Ya., et al.: Space-Time Signal Processing, ed. by Kremer, I.Ya. Radio i Svyaz, M. (1984). (in Russian)
2. Komarovich, V.F., Nikitchenko, V.V.: Methods of Spatial Signal Processing. VAS, L. (1989). (in Russian)
3. Avdeyenko, G.L., Fedorov, V.I., Yakornov, Ye.A.: Determination of the radiation source location based on the electromagnetic wave's front curvature. Radioelectron. Commun. Syst. **51**(3), 115–121 (2008). <https://doi.org/10.3103/S0735272708030011>
4. Avdeyenko, G.L., Lipchevskaya, I.L., Yakornov, E.A.: Phase systems of determining coordinates of radiation source with harmonic signal in Fresnel zone. Radioelectron. Commun. Syst. **55**(2), 65–74 (2012). <https://doi.org/10.3103/S0735272712020021>
5. Avdeyenko, G.L., Ilchenko, M.Yu., Yakornov, Ye.A., et al.: UA Patent No. 64696, Byull. Izobret., No. 21 (2011)
6. Vishnevsky, V., Shahnovich, I., Frolov, S.: Radio-relay communication lines in the millimeter range, new horizons of speeds. Elektronika NTB **1**(107), 90–97 (2011). (in Russian)
7. Ilchenko, M.Ye., Narytnyk, T.M., Fisun, A.I., Belous, O.I.: Terahertz range telecommunication systems. Telecommun. Radio Eng. **70**(16), 1477–1487 (2011). <https://doi.org/10.1615/telecomradeng.v70.i16.60>
8. Ilchenko, M.Ye., Kuzmin, S.Ye., Narytnyk, T.M., Fisun, A.I., Belous, O.I., Radzikhovskiy, V.N.: Transceiver for 130–134 GHz band digital radio relay system. Telecommun. Radio Eng. **72**(17), 1623–1638 (2013). <https://doi.org/10.1615/telecomradeng.v72.i17.70>
9. Yakornov, Ye.A., Kolomytsev, M.A., Avdieienko, H.L., Lavrinenko, O.Yu.: Theoretical analysis of possibility of electromagnetic wave phase front curvature physical phenomenon application in stationary microwave radio systems. Visnyk NTUU KPI Seriiia - Radiotekhnika Radioaparotobuduvannia **48**, 84–96 (2012). <https://doi.org/10.20535/RADAP.2012.48.97-106>
10. Avdeyenko, G.L., Kolomytsev, M.A., Yakornov, Ye.A.: Efficiency of spatial signal processing in wireless communications. Telecommun. Sci. **3**(2), 5–13 (2012)
11. Avdeyenko, G.L., Yakornov, Ye.A.: Optimal spatial processing effectiveness indexes calculation for fixed wireless communication system radiolink. Visnyk NTUU KPI Seriiia - Radiotekhnika Radioaparotobuduvannia **52**, 92–101 (2013). <https://doi.org/10.20535/RADAP.2013.52.92-101>
12. Ilchenko, M.Ye., Yakornov, Ye.A., Avdeyenko, G.L., et al.: UA Patent No. 104240, Byull. Izobret., No. 2 (2016)
13. Ilchenko, M.Ye., Yakornov, Ye.A., Avdeyenko, G.L., et al.: UA Patent No. 104241, Byull. Izobret., No. 2 (2016)



Fiber-Optic and Laser Sensors-Goniometers

S. Ivanov  

National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Peremoga Avenue 37, Kyiv 03056, Ukraine
marinex@inbox.ru

Abstract. Fiber-optic and laser sensors – goniometers are sensors for measurement of rotation angle of an object. Those sensors can be used in many areas of science and technology, including telecommunications (measurement of rotation angle of an antenna).

Creation of this class of devices became possible only with the development and improvement of the elemental basis of quantum electronics. The principle of operation of optical goniometers is based on the vortical effect of Sagnac.

The interest of foreign firms in optical goniometers is based on their potential applications in many fields of science and technology. In some cases these devices can completely replace difficult expensive electromechanical decision. The Institute of Telecommunication Systems has developed some types of such sensors based on the original engineering solutions.

Keywords: Fiber-optic goniometer · Laser goniometer · Rotation angle meter

1 Introduction

In many fields of science and technology, including telecommunications, precise measurement of angle is an important task. In order to measure angle, specialized instruments – goniometers are used. One of the promising ways of development is the optic goniometry based on the Sagnac effect.

Optical goniometers have a number of advantages:

- (1) high precision of angle measurement;
- (2) small size, mass and low energy consumption (fiber-optic goniometers);
- (3) high reliability, as there are no moving parts.

Problems, related to development of laser goniometers, are discussed, for example, in works [1, 2], of fiber-optic ones, in works [3–8]. The Institute of Telecommunications Systems, based on the published data and its own research, conducts research in the field of optic goniometer development and development of such instruments.

In this section, achievements of the Institute of Telecommunications Systems in the development of such sensors are presented.

2 Four-Frequency Laser Goniometer

Classic laser goniometer is a ring gas laser, in which two counter-rotating waves with linear polarization are generated. After passing through the contour in opposite directions, waves get out of the contour and are fed to the photodetector, where interference takes place. Goniometer's photodetector has two photosensitive areas, which are offset one to another by the quarter of interference pattern period. During rotation of an object, on which goniometer is installed, that pattern moves with the speed, proportional to difference of frequencies of counter-rotating waves, and as a result two signals, shifted by phase by a quarter of period, emerge at the photodetector output. Output signal of such a sensor is the number of pulses, registered during a given time, which is proportional to the angle of object rotation; direction of object's rotation is determined using the phase shift between signals from photodetector's areas.

Because of the "frequency lock-in" phenomenon (synchronization of frequencies of counter-rotating waves), laser goniometer has a significant drawback, namely a dead zone at low object's rotation velocities; that significantly decreases precision of the rotation angle measurement. In order to solve that problem usually a so-called vibration suspension is used: sensor's monoblock is installed on the mechanical suspension, with the help of which said monoblock oscillates with a small magnitude about sensor's sensitivity axis (forced dithering). The drawback of such a decision is an increased complexity of sensor's design and appearance of the moving parts; that decreases sensor's reliability.

Four-frequency laser goniometer allows one to eliminate dead zones without use of any moving parts (vibration suspension), and also allows one to improve precision and reliability of the sensor. To that end, four-frequency operation mode with the use of circularly polarized modes is proposed to be used in that goniometer.

In the contour of a goniometer, two pairs of counter-rotating waves with orthogonal polarization (one pair has right polarization, another one – left polarization) are simultaneously excited. As a result, two virtual laser goniometers work in one contour. In order to get both those goniometers out of the lock-in zone, Faraday cell is used. As polarizations for said virtual goniometers are different, Faraday element produces in each of them the frequency shift; those shifts are equal by value but have different signs. As a result, in the four-frequency sensor's output signal, initial frequency shifts, induced by the Faraday cell, are subtracted and compensate each other, and frequency shifts, caused by rotation (desired signal) are added and amplify each other. Such a scheme allows one (in a linear approximation) to eliminate errors, induced by the Faraday element, and also to decrease demands to the precision of the frequency bias, decrease influence of the external magnetic field. Moreover, one can use permanent magnets in order to use the initial frequency shift instead of the alternating frequency bias; that allows one to eliminate noises, induced by such a frequency bias and improve linearity of the output characteristics. One should note, that use of the circular polarization itself decreases connection of the counter-rotating waves through the backscattering, and as a result, dead zone becomes narrower.

Other advantages of the four-frequency laser goniometer are:

- increase of the sensor's sensitivity by two times;
- decrease of noises because of the mutual error compensation;
- faster warm-up and lower warm-up time;
- lower sensitivity of the sensor to ambient conditions.

Design of such a sensor is developed in the research department, along with model of its errors, electronics and firmware. Overall view of the device is given in Fig. 1.

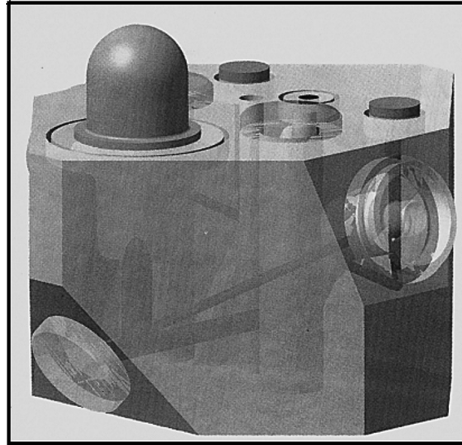


Fig. 1. Four-frequency laser goniometer

Zero bias of such a goniometer is $0,005^\circ/h$.

3 Block of Laser Goniometers for Aircraft Inertial Navigation System

Laser goniometer measures angle of rotation of object, on which it's installed relative to the inertial space. That allows one, with the use of a block of three such goniometers, to determine orientation of object, on which such a block is installed, and to use those data in inertial navigation system of a civil aircraft.

In Fig. 2, scheme for modeling of such a block in Ansys environment is shown.

As one can see from Fig. 2, block of goniometers consists of goniometers 1, 2 and 3 themselves and also a frame 4, on which they're installed. Shock absorbers 5 are intended to isolate devices from the vibration and shocks of the aircraft (e.g., during take-off and landing). In addition, in order to ensure that center of gravity of the block is in the predetermined point, one has to use counterweights (not shown in Fig. 2).

The fundamental problem during use of such a block of three goniometers is the fact, that because of interaction of their vibration suspensions, a conical movement of instruments' sensitivity axes arises; as a result, precision of orientation determination decreases. Paper [8], published by the department's workers, is devoted to creation of the model of interaction of vibration suspensions.

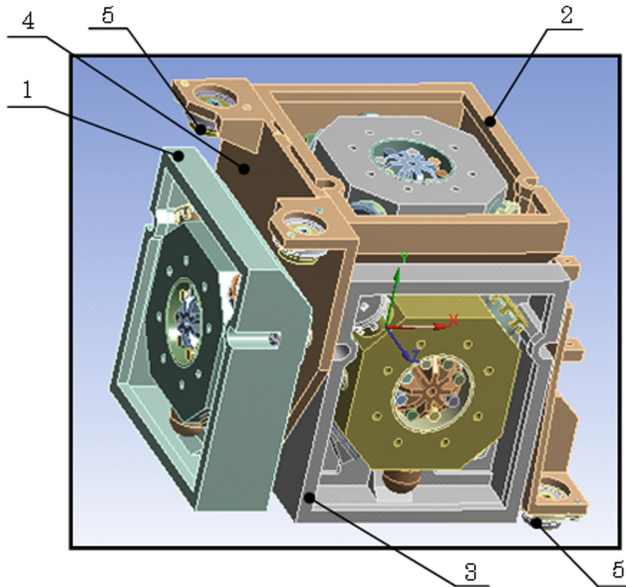


Fig. 2. Block of laser goniometers of aircraft inertial navigation system (scheme)

Authors of [8] developed model of displacements that take place in the block of strapdown inertial navigation system's (SINS) elements because of interaction of vibration suspensions of three laser goniometers and identified parameters of that model. The approach used allowed one to create analytical model, which describes internal oscillation processes, and to explore each of them independently of others as a cause of degradation of precision of block of sensors operation.

The model obtained significantly decreases time needed for modeling, which will allow one to make study of SINS operation significantly easier when varying its parameters, including parameters of frame and vibration suspension control signal, and also to develop effective control in order to minimize errors, caused by the interaction of sensors' vibration suspensions.

On the base of that study, research department has developed and produced electronics and firmware of the block of three goniometers. In particular, vibration suspension control algorithm was developed that allows one to improve zero drift of the sensor on a non-moving base up to $0,001^\circ/\text{h}$.

Zero bias of such a block is $0,01^\circ/\text{h}$; that value is sufficient for use in the civil aviation.

4 Three-Axis Laser Goniometer

Three-axis laser goniometer is intended to measure angles of object rotation in the orientation system. Advantages of such goniometer compared to the traditional block of three sensors are:

- lower mass, as there're no need in frame and heavy counterweights;
- smaller size;
- lower energy consumption, compared to the block of three goniometers;
- there's only one vibration suspension in the sensor, therefore problems with the conical movement of sensitivity axes and interaction of vibration suspensions are ruled out.

Mock-up of the monoblock of such a goniometer is shown in Fig. 3; channels of sensor's parts that correspond to different sensitivity axes are marked with different colors. The distinctive feature of such a three-axis goniometer is the fact, that, in contrast to classic and four-frequency goniometers, contours, sensitive to rotation, are passive. That means that laser radiation is not generated in closed contours directly, but is fed into all those contours from the same external laser. That allows one not only to decrease sensor's energy consumption (one gas laser instead of three), but also to eliminate dead zone, as in passive contours "frequency lock-in" phenomenon is absent. Similar principle of operation is used in fiber-optic goniometers, which are discussed below.

One should note, that three-axis goniometer also has some drawbacks, namely:

- monoblock is hard to manufacture;
- complex algorithm of measured data processing is needed, as axes of sensitivity of three parts of the sensor are not perpendicular to each other.

Sensor design is developed by the research department; technical documentation needed for manufacturing of the sensor is also prepared.

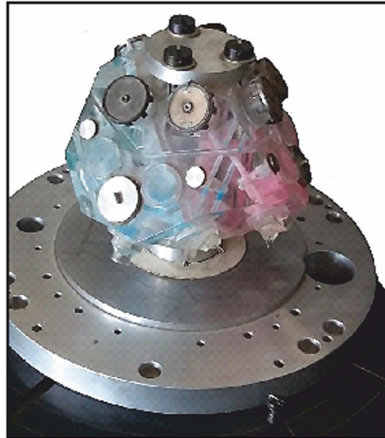


Fig. 3. Monoblock of a three-laser goniometer (mock-up)

Zero bias of a three-laser goniometer is $0,02^\circ/\text{h}$.

5 Interferometric Fiber-Optic Goniometers

In spite of the fact, that interferometric fiber-optic goniometer is similar to laser one in principle of its operation, there're significant differences between those sensors. In fiber-optic goniometers, sensitive element is always a passive closed loop of an optic fiber; radiation into that loop is fed from an external source. Significant advantage of such technical decision is an absence of the "frequency lock-in" phenomenon, which is specific to the laser goniometer. As sensitivity of the fiber-optic goniometer is directly proportional to the loop area, in order to increase sensitivity a multiturn coil is used, at that length of the fiber wound is from 100 m up to some km in high-precision sensors (for comparison: in a laser goniometer, perimeter of the laser contour is tens of cm).

As a light source in the interferometric not laser, but a superluminescent diode (SLD) is used. That's due to the fact, that laser's coherence length is some meters, and SLD's one is some parts of mm. Because of that, in the output signal noises, that emerge as a result of Rayleigh backscattering, polarization noises and other optical noises make contribution into the interference pattern in the form of a constant term, which can be eliminated during signal processing (use of the laser leads to the emergence of the noise which is very hard to filter out). In order to provide the given range of rotation velocities, in which goniometer can measure rotation angle, SLD coherence length is selected to be comparable with the maximal difference of optical lengths of paths of light propagation through the loop in opposite directions, which is proportional to the object's rotation speed.

Interferometric fiber-optic goniometers operate using two schemes: direct measurement (open-loop) and compensational (closed-loop) ones. Scheme of the open-loop sensor is shown in Fig. 4.

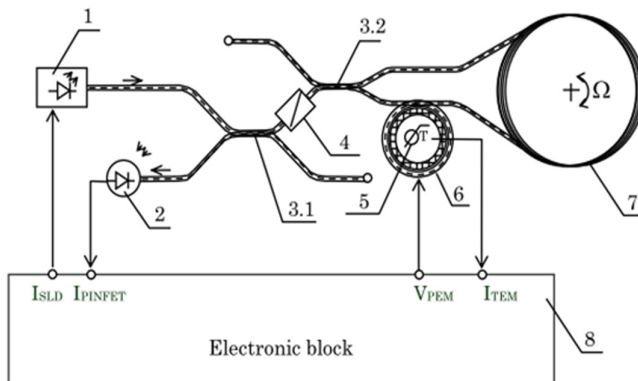


Fig. 4. Scheme of the open-loop fiber-optic goniometer: 1 – superluminescent laser diode (SLD); 2 – photodetector; 3 – fiber coupler; 4 – polarizer; 5 – thermosensor; 6 – piezoelectric phase modulator; 7 – fiber loop; 8 – electronic block

Light source of such a goniometer is SLD 1, and an output radiation is registered by spectrum- and polarization-insensitive photodetector 2, which photocurrent is proportional to the radiation intensity.

Fiber couplers 3 are used to mix/divide light beams that propagate in fibers. First coupler is intended to transfer the light, which passed through the optic contour, to the photodetector 2. Radiation of the source before feeding into the loop is divided by the second coupler into two beams, which are fed into the loop in opposite directions. After the propagation through the contour, light beams are mixed on the second coupler and interfere on the photodetector.

Polarizer 4 decreases zero shift, caused by irregularities of the optic contour; it determines the level of parasitic signal at the output.

Phase shift between counter-rotating waves is introduced by piezoelectric modulator 6, which is placed on the one fiber loop ends. Sinusoidal voltage at the frequency of radial resonance of the modulator is fed to it, at that phase of modulation of one wave differs from the phase of modulation of the opposite one by $2\pi v_m \tau$ (here and hereinafter τ is the time of propagation of the electromagnetic wave in the loop).

Electronic block controls sensor's phase modulator, SLD power current, and also provides acquisition of the electric signal from photodetector, its following processing and output of sensor's rotation angle information in the digital form. The essence of the open-loop goniometer signal processing algorithm is measurement of Sagnac angle that emerges as a result of sensor rotation and output of object's rotation angle estimation based on that Sagnac angle.

Open-loop scheme is simple and easy to implement, doesn't require high computational power of the electronic block because of simple processing algorithm.

In the course of the development of interferometric fiber-optic goniometer, built using open-loop scheme, influence of its elements on measurement precision was analyzed [9]. It was shown during that study, that polarizer and modulator have the most significant influence on instrument precision. One should pay particular attention to modulator improvement during sensor design, as error, induced by its imperfection, is multiplicative and can take inadmissibly high values. Studies also showed that to decrease measurement error one should select SLD with stable polarization and radiation ellipticity parameters.

Drawbacks of the open-loop scheme are:

- narrow range of angular velocities, in which goniometer can operate;
- low resolution (the rougher sensor is, the wider range of angular velocities is and vice versa);
- nonlinearity of scale coefficient out of some range of object's angular velocities.

Because of that, that scheme can be used only on low-precision angle sensors. In order to create medium-precision sensor, scheme with the closed control loop (or, simply, closed-loop scheme) is used as a rule.

Closed-loop goniometer is a compensational instrument, i.e. principle of its operation is based not on the measurement of the Sagnac angle, emerged as a result of rotation, but compensation of its influence on the output signal using modulation of SLD radiation. As a result, the time sensor operates in the vicinity of the operating

points, selected on the linear part of its output characteristic. That gives one some advantages:

- high stability of the scale factor, and, as a result high linearity of the sensor's output characteristic;
- significant (by some times) widening of the range of angular velocities, in which goniometer can operate, as when Sagnac angle is compensated, restriction in the forms of non-linearity of the characteristic and crossing of 2π value by Sagnac angle;
- high resolution, as there are no limitations on the range of angular velocities and one can significantly increase instrument's sensitivity.

Drawbacks of the closed-loop scheme are:

- relative complexity of data processing algorithm;
- one needs higher computational power compared to open-loop goniometer;
- if at the moment of sensor's power-on Sagnac angle has already crossed 2π value, it's impossible to detect without an additional sensor, has Sagnac angle already crossed 2π value or not, and goniometer becomes inoperable.

Scheme of the closed-loop fiber optic goniometer is shown in Fig. 5.

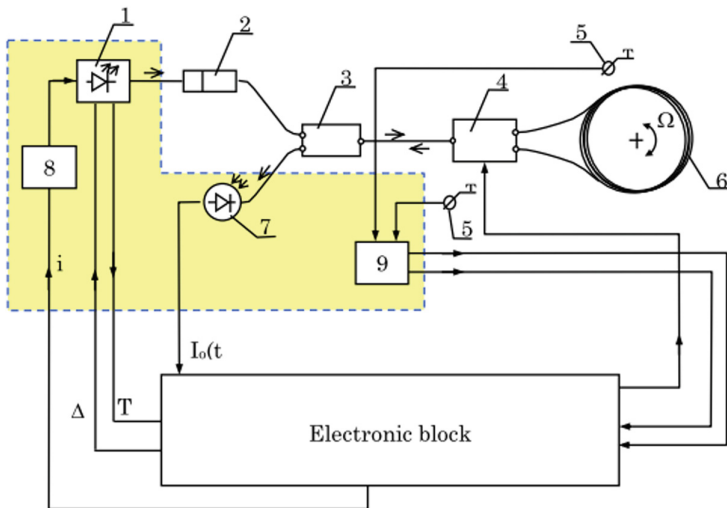


Fig. 5. Scheme of the closed-loop fiber optic goniometer: 1 – superluminescent diode (SLD), 2 – Lyot depolarizer, 3 – coupler, 4 – integral-optic modulator, 5 – thermosensor, 6 – fiber loop, 7 – photodetector, 8 – amplifier of SLD operating current control system, 9 – convertor of signals of thermosensors

Light source in the goniometer is SLD 1, its radiation through Lyot depolarizer 2, which is intended to decrease sensor's zero bias, and the coupler 3 is fed to the integral-optic modulator 4.

Integral-optic modulator provides three functions:

- (1) polarization of light in order to decrease zero bias instability because of polarization nonreciprocity;
- (2) dividing of light waves into ones with equal power that propagate in the fiber loop clockwise and counterclockwise, with the following recombination on the coupler;
- (3) introduction of the phase shift between counter-rotating waves using electro-optic phase modulator.

Sensor’s output signal is registered by spectrum- and polarization-insensitive photodetector 2, which photocurrent is proportional to the radiation intensity.

Electronic block provides sensor powering, processing photodetector signal and control of the phase modulator. Electronic block also provides (through the amplifier 8) SLD current control, stabilization of its output power and temperature.

Control signal from the electronic block also corrects modulator operation such a way, that it compensates phase shift, that emerges as a result of sensor rotation. If full compensation is achieved, difference of levels of positive and negative half-waves of the voltage, coming from the photosensor to the ADC of the control block, is equal to zero. Correction of modes of operation of the fiber-optic goniometer is done based on the thermal sensors 5 readings.

For operation of the closed-loop goniometer, modulator control method is important [10]. In the classic variant, voltage at the modulator is a sequence of rectangular pulses (see Fig. 6), each of those at $U_{\pi/2}$ voltage adds phase shift $\pi/2$ to the wave.

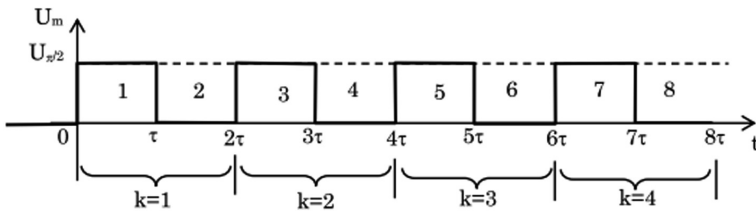


Fig. 6. Classic modulator control

Such modulator control method allows one to measure Sagnac angle during time of two propagations through the contour (2τ). The drawback of the method is restriction of the range of angular velocities, in which goniometer can operate, by value

$$|\Omega| < \frac{\pi}{2K_0} \tag{1}$$

where K_0 is the optic scale factor of a goniometer.

In work [9], modulator control using the complex periodic signal, shown in Fig. 7, is described.

Use of the control method, that uses complex periodic signal, allowed one, at the same interferometer configuration, to improve dynamic range of angular velocities compared to the classic method, up to the value

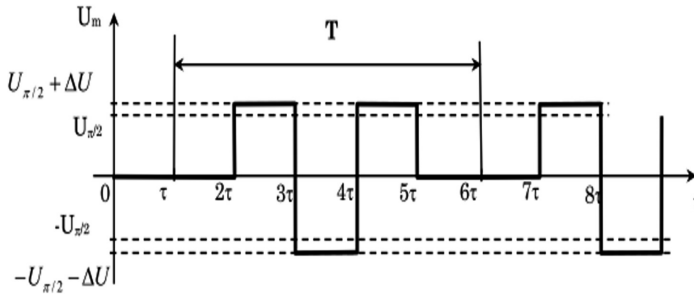


Fig. 7. Modulator control using the complex periodic signal

$$|\Omega| < \frac{2\pi}{K_0} \tag{2}$$

Moreover, use of the control with complex periodic signal allowed one to decrease influence of the thermal and time drifts of radiation source intensity on the detection of the phase shift. The drawback of such control is the decrease of sensor’s sensitivity because of the increase of the Sagnac angle measurement time from 2τ to 5τ (period of the modulation signal).

The most important error source in the interferometric goniometer is the Shupe effect. The cause of the effect is the fact, that when temperature changes, coil is heated gradually (e.g., from external layers to internal ones, if external temperature is rising). As a result, during propagation through the loop, counter-rotating waves pass the same sections of the fiber at different times and at different values of the temperature, i.e. at different lengths of the fiber sections (because of thermal expansion) and different refraction indices. That causes Shupe effect: during propagation through the loop difference of phases of counter-rotating waves is accumulated, that difference is absolutely not related to sensor’s rotation, cannot be extracted from output signal and is very hard to compensate for.

In order to eliminate that effect, except the design means (thermal insulation) and thermostating one uses special winding types, in particular dipole and quadrupole ones. Analysis of effectiveness of different winding types is made in [11] based on the mathematical modeling of the Shupe effect in goniometer with different winding types. As an example. In Fig. 8, plot of zero bias vs. time under the simultaneous influence of radial and axial gradients on the coil with different winding types wound on the spool (carcass), and in Fig. 9 – plot of zero bias vs. time under the same influence on the coil with padding (or frameless coil) with different winding types.

One can conclude from the analysis, made in [11], that in general, from the point of view of ensuring sensor’s precision, use of quadrupole windings is optimal. For low-precision sensors, one can use one-pole centered winding as the most simple and cheap in production, but only if good thermal and vibration insulation of the coil is provided.

It’s also evident from Figs. 8 and 9, that for simple winding, bifilar winding and crossover-free winding zero bias for coil with paddings (or frameless coil) is decreased by two times; for other winding types decrease is insignificant.

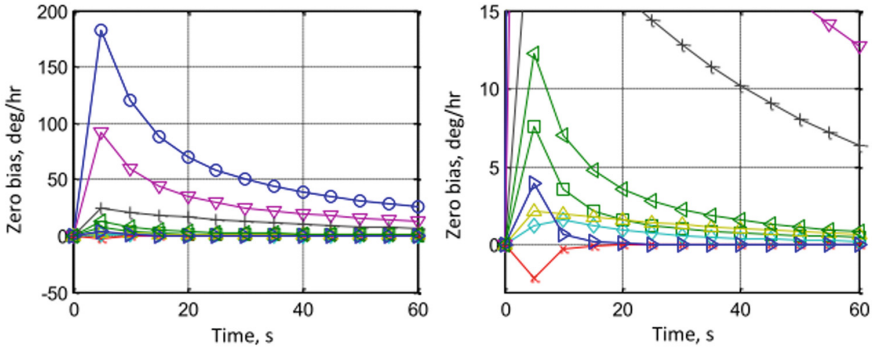


Fig. 8. Zero bias under the influence of radial and axial gradient on the coil with winding:
 —○— simple (one-pole), —□— dipole, —×— quadrupole, —◇— modified quadrupole,
 —▽— bifilar, —△— centered quadrupole, —+— centered one-pole,
 —▶— octupole, —◀— crossover-free

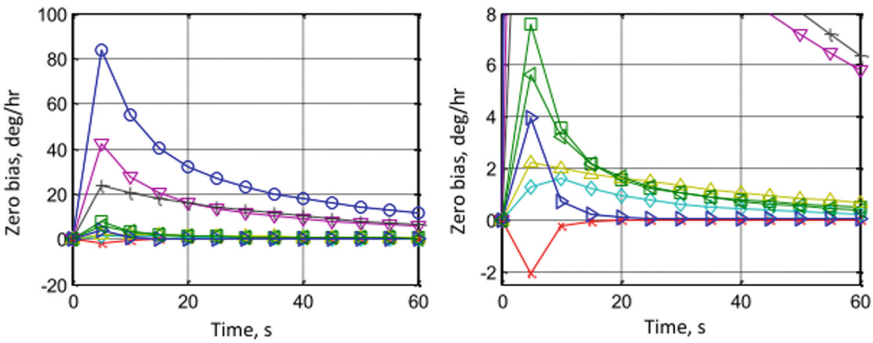


Fig. 9. Zero bias under the influence of radial and axial gradient on the coil with padding (or frameless coil) with winding:
 —○— simple (one-pole), —□— dipole, —×— quadrupole, —◇— modified quadrupole,
 —▽— bifilar, —△— centered quadrupole, —+— centered one-pole,
 —▶— octupole, —◀— crossover-free.

Thus, use of paddings or frameless coil is justified for production of low-precision goniometers (simple or bifilar winding) or if crossover-free winding is used. For all other windings, symmetrical relative to influence of radial and axial gradients, use of paddings leads only to insignificant decrease of the zero shift, thus use of complex winding plus special paddings combination is justified only for high-precision instruments.

Low-precision (see Fig. 10a), medium precision (Fig. 10b) and three-axial (Fig. 10c) fiber-optic goniometers are developed by the research department. Three-axis sensor contains three independent optic contours, light into those contours is fed from the same source; sensor is created using open-loop scheme.

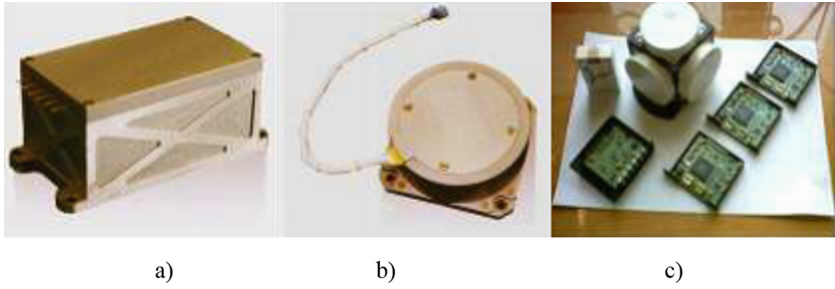


Fig. 10. Goniometers, developed by the research department: (a) low precision; (b) medium precision; (c) three-axial.

Zero bias of the medium-precision goniometer is not worse than $0,05^\circ/h$, and for other sensors is $0,3 \dots 0,8^\circ/h$.

6 Resonant Fiber-Optic Goniometers

Digital resonant fiber-optic goniometer is based on the Sagnac effect as well as interferometric one. Difference between those sensors lies in the principle of rotation registration, sensor design and signal processing methods.

Sensitive element of the sensor is the resonator – closed fiber loop, into which laser radiation is fed; part of that radiation after multiple turns in the loop is get of it and fed to the photodetector. Principle of operation of the resonant goniometer is as follows: when resonator rotates, its resonant frequencies for waves, propagating within in opposite directions, are shifted one relative to another by value, proportional to the rotation velocity. If values of resonant frequencies of such a loop are known, one can estimate angular velocity and rotation angle of an object.

In Fig. 11 resonant curves during goniometer rotation for reflection-type resonator (a) and transmission-type resonator (b) are shown.

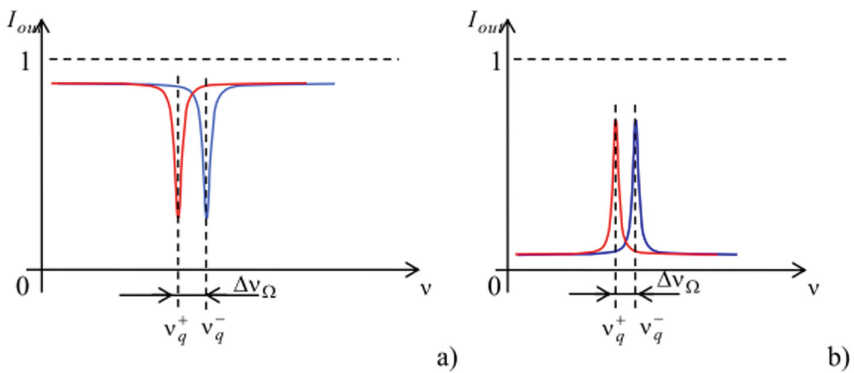


Fig. 11. Resonant curves during goniometer rotation: (a) for reflection-type resonator; (b) for transmission-type resonator

In the reflection-type resonator, light is fed into the resonator and output out of it in the same point, waves compensate each other, and when radiation frequency is equal to the resonant one, signal intensity at the contour output is minimal (see Fig. 11a). In the transmission-type resonator, radiation is fed into the loop and output out of it in different points, therefore when radiation frequency is equal to the resonant one, signal intensity at the contour output is maximal (see Fig. 11b). Blue curve in Fig. 11 is plotted for the wave that propagates in the resonant loop clockwise, and red- for the counter-clockwise one; one can see, that if rotation is present, waves are shifted from some medium position.

In order to measure difference of the resonant frequencies of the goniometer contour for counter-rotating waves, one has to adjust laser radiation frequency for each direction separately and hold output signal on the minimal or maximal level correspondingly. At that, as measurements are made in the vicinity of antiresonance (or resonance, depending on the scheme type – reflection-type or transmission-type one), in order to obtain acceptable precision linewidth of the radiation source is very important. SLDs are unfit for use in the resonant fiber-optic goniometer because of wide radiation spectrum (up to tens of nm). Laser diodes, Distributed Feedback Lasers and other highly coherent light sources with radiation spectrum linewidth of 1–100 kHz are used as a light source in such a goniometer. At that, the better resonant loop is (the higher is its finesse, and, therefore, potential instrument sensitivity), the narrower light source spectrum should be. Scheme of the reflection-type goniometer is shown in Fig. 12.

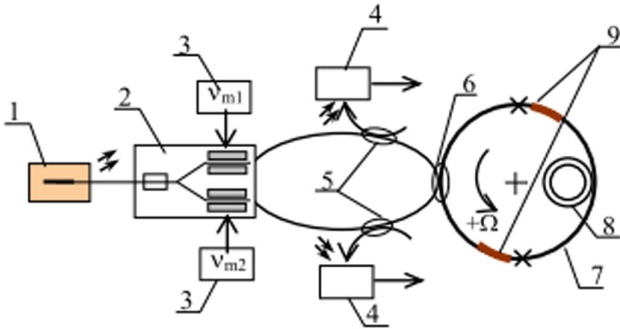


Fig. 12. Optical scheme of the reflection-type goniometer: 1 – fiber laser; 2 – integral-optic modulator; 3 – special function generator; 4 – photodetector; 5–50% coupler; 6 – directional coupler; 7 – resonant fiber loop; 8 – piezocorrector; 9 – polarizers

Radiation of the fiber laser 1 is fed into the integral-optic modulator 2, where is divided into two beams with equal intensities; phases of waves of those beams are modulated at frequencies ν_{m1} and ν_{m2} . After propagation through modulator 2 and couplers 5, beams by the means of directional coupler 6 are fed into the resonant fiber loop 7, made of the photonic-crystal fiber. After multiple turns in opposite directions, beams re output out of the contour through the same directional coupler 6 and through

couplers 5 are fed to photodetectors 4. Signals of photodetectors 4 are used to adjust radiation frequency of the laser 1 to the contour's resonance and estimation of the rotation angle. In order to decrease polarization noise, in two reciprocal points of the contour 7 fiber is spliced and joined with the rotation of anisotropy axes by 90° , and two polarizers 9 are installed. Piezocorrector 8 is used to minimize thermal influence on the goniometer.

Scheme of the transmission-type goniometer is shown in Fig. 13.

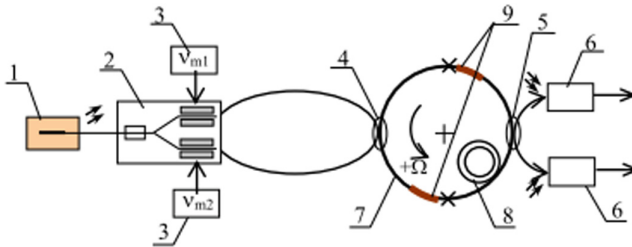


Fig. 13. Optical scheme of the transmission-type goniometer: 1 – fiber laser; 2 – integral-optic modulator; 3 – special function generator; 4 – directional coupler; 5 – directional coupler; 6 – photodetector; 7 – resonant fiber loop; 8 – piezocorrector; 9 – polarizers

Operation of that scheme is similar to the scheme, shown in Fig. 12, save that the light is output form another directional coupler.

Advantages of the resonant goniometers over traditional interferometric ones are small length of the fiber loop (resonator) – from some cm up to 10–12 m; that allows one in the future to decrease sensor size down to the level of microoptic one, with passive optic resonator on the silicon chip.

As drawback, one should point the a relatively complex control system has to be used; that system should track resonant frequencies of counter-rotating waves, which significantly depend not only on sensor's rotation, but also on temperature.

Studies of that sensor type are now active, and one cannot speak about the final result. However, studies made allow one to make some conclusions.

According to results of studies, main problems of resonant goniometer development are:

1. High losses in the optical resonant contour, that decrease its quality factor, finesse, and, as a result, instrument sensitivity. In order to solve that problem, one has to use custom fibers and optic components.
2. High losses on optic components outside the resonator, which significantly decrease instrument sensitivity and call to increase of power of the laser.
3. In spite of the fact, that Shupe effect in resonant goniometers is negligible because of short fiber loop, sensitivity of the contour's resonant frequency to temperature because of thermal expansion is one of key factors, that decrease precision of angle measurements.

4. Highly coherent radiation sources, which radiation frequencies can be adjusted, are fit to the use in the goniometer. One of the variants of solution of the problem is the use of serrodyne modulation of a radiation.
5. Device calls for the use of a complex control system, which, among other functions, has to control radiation source frequency, track changes contour's in resonant frequencies, changes of fiber loop temperature, etc.
6. Excepts temperature, backscattering and polarization has a significant influence on the sensor's precision.

Problems, described above, are now solved by the research department of RDI of telecommunications.

Use of different modulation types in a resonant goniometer is an important problem.

As a rule, sinusoidal, sawtooth, triangular and hybrid serrodyne (see Fig. 14) modulation is used in resonant goniometers.

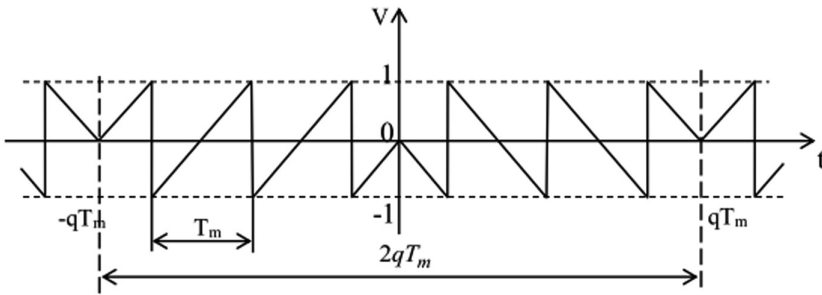


Fig. 14. Hybrid serrodyne modulation

Use of a sawtooth modulation allows one to provide laser frequency shift; triangular and hybrid serrodyne modulation allow one to simplify estimation of loop's resonant frequencies for counter-rotating waves and data processing algorithm.

Modeling of the device with different modulation types showed that use of sinusoidal and hybrid serrodyne modulation calls for precise hold of modulation index, and the most effective form from the point of view of measurement system implementation is the triangular modulation.

Also as a result of analysis of state-of-the-art trends and own studies, it can be stated, that use of the fully digital modulation and digital control systems in resonant goniometers is very promising.

Now research the Institute of Telecommunications Systems develops resonant fiber-optic goniometer based on the results of the studies obtained. Zero bias of such a goniometer will be $0,1^\circ/h$.

7 Conclusions



As one can see from the stated above, the Institute of Telecommunications Systems developed some types of fiber-optic goniometers, including promising four-frequency laser ones, and interferometric fiber-optic goniometers, based on own research. Now the Institute of Telecommunications Systems develops resonant fiber-optic goniometer, which is promising from the point of view of sensor minimization.

References

1. Aronowitz, F.: Fundamentals of the ring laser gyro. *Opt. Gyros Appl.* **339**, 3.1–3.45 (1999)
2. Lahham, J.I.: Tuned support structure for structure-borne noise reduction of inertial navigation with dithered ring laser gyros (RLG). In: Lahham, J.I., Wigent, D.J., Coleman, A. L. (eds.) *Position, Location and Navigation Symposium, PLANS*, San Diego, USA, pp. 419–428 (2000). <https://doi.org/10.1109/plans.2000.838334>
3. Yin, Sh, Ruffin, P.B., Yu, F.T.S. (eds.): *Fiber Optic Sensors*, 2nd edn, p. 494. CRC Press, Boca Raton (2008)
4. Chen, X., Shen, C.: Study on temperature error processing technique for fiber optic gyroscope. *Optik – Int. J. Light Electron Opt.* **124**(9), 784–792 (2013). <https://doi.org/10.1016/j.ijleo.2012.02.008>
5. Jia, M., Yang, G.: Research of optical fiber coil winding model based on large-deformation theory of elasticity and its application. *Chin. J. Aeronaut.* **24**, 640–647 (2011)
6. Zhou, K., Hu, K., Dong, F.: Single-mode fiber gyroscope with three depolarizers. *Optik – Int. J. Light Electron Opt.* **125**(2), 781–784 (2014). <https://doi.org/10.1016/j.ijleo.2013.07.081>
7. Medjadba, H., Lecler, S., Simohamed, L.M., Fontaine, J., Meyrueis, P.: Investigation of mode coupling effects on sensitivity and bias of a multimode fiber loop interferometer: application to an optimal design of a multimode fiber gyroscope. *Opt. Fiber Technol.* **17**(1), 50–58 (2011). <https://doi.org/10.1016/j.yofte.2010.10.004>
8. Ivanov, S.V., Volovyk, V., Slabukhin, I.S.: Rozrobka modeli vzaiemovplyvu vibropidvisiv lazernykh hiroskopi v BINS (Development of model of interaction of laser gyros' vibration suspensions in the strapdown inertial system). *East.-Eur. J. Enterp. Technol.* **3**(7), 42–47 (2015)
9. Ivanov, S.V., Ilkiv, M.I.: Metod keruvannia volokonnyim interferometrom iz zamknytim konturom zvorotnoho zviazku (Control method of fiber interferometer with closed feedback loop). *Informatsiini systemy, mekhanika ta keruvannia: naukovo-tekhnicnyi zbirnyk* (9), 113–124 (2013). Refs.: 2 titles
10. Ivanov, S.V.: Vplyv parametriv elementiv volokonno-optychnoho hiroskopa z vidkrytoiu petleiu zvorotnoho zviazku na tochnist vymiriuvannia (Influence of parameters of open-loop fiber-optic gyro's components on measurement precision). *East.-Eur. J. Enterp. Technol.* **1** (9), 16–24 (2016)
11. Ivanov, S.V.: Porivnialnyi analiz efektyvnosti vydiv namotky volokna chutlyvoho elementa volokonno-optychnoho hiroskopa v umovakh zminy temperatury (Comparative analysis of the efficiency of fiber winding types used in the fiber-optic gyro sensor coil under temperature variation conditions). *Naukovi visti NTUU «KPI»: mizhnarodnyi naukovo-tekhnicnyi zhurnal* **1**(105), 99–107 (2016)



Bridge Equivalent Circuits for Microwave Filters and Fano Resonance

M. Ilchenko^(✉)  and A. Zhivkov^(✉) 

National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Peremohy Avenue 37, Kyiv 03056, Ukraine
{ilch, a. zhivkov}@kpi.ua

Abstract. The analytical dependences for analysis of microwave structures based on different types of resonators located in parallel communication channels are obtained. These dependencies are suitable for describing many known oscillatory processes. A comprehensive analysis of electrodynamic metamaterial cells is carried out on the basis of these equations. It was shown that obtained analytical expressions describe the characteristics of bridge 4-poles on lumped elements. Such 4-pole networks have an extremely high group delay time and complete signal rejection possibility. These properties characterize metamaterials and some processes described by the Fano resonance. By contrast 4-poles based on coupled resonators, bridge 4-pole are built on the basis of two independent oscillations in different branches. The purpose of this chapter is new models of bridge 4-poles with different Q-factors of resonators. These models simplify modeling of attenuation poles in microwave filters and processes described by Fano resonance.

Keywords: Bridge equivalent circuits · Metamaterials · Fano resonance

1 Microwave Filters with Resonators in Parallel Channels. Equations and Programs for Calculating Filters Based on Mutually Detuned Resonator in Parallel Channels

Microwave structures with not interconnected resonators located in parallel channels (RPC) were analyzed in [1]. We will differ them from structures with mutually connected resonators (MCR). In [2] it was shown that structures with RPC are elementary cells of electrodynamic metamaterials. Analysis of elementary cells of metamaterials using programs based on analytical expressions [1] allowed for detecting some interesting properties [3]. The transmission \mathbf{T} and reflection \mathbf{R} coefficients for two-resonance bandpass filter (BPF) and bandstop filters (BSF) based on the RPC, taking into account different Q-factors of resonators, are given below:

BPF:

$$\mathbf{T} = K_2/(1 + K_2) - K_1/(1 + K_1) \quad (1)$$

$$\mathbf{R} = 1 - K_1/(1 + K_1) - K_2/(1 + K_2) \quad (2)$$

where K_1, K_2 —coupling coefficients of resonators of transmission line:

$$K_1 = K_1/[1 + j(\xi + a)] \quad (3)$$

$$K_2 = K_2/[1 + j(\xi - a) * b] \quad (4)$$

$$K_1 = Q_{01}/Q_{1l} \quad (5)$$

$$K_2 = Q_{02}/Q_{2l} \quad (6)$$

where Q_{0i} is unloaded quality factor of “i” resonators, Q_{li} is loaded quality factor of “i” resonators; ξ is the generalised detuning of frequency regarding to the central frequency f_0 of the filter, and a is the generalised detuning of “magnetic” f_m and “electric” f_e frequencies regarding to f_0

$$f_0 = (f_m + f_e)/2 \quad (7)$$

$$a = 2(f_e - f_m)/(f_e + f_m) \times \sqrt{Q_{01} * Q_{02}} \quad (8)$$

$$\xi = 2(f - f_0)/(f + f_0) \times \sqrt{Q_{01} * Q_{02}} \quad (9)$$

$$b = \sqrt{Q_{01}/Q_{02}} \quad (10)$$

It is important to note that the use of terms “electric” and “magnetic” in relation to oscillation types is correct if resonators are modeled by electric and magnetic dipoles. These oscillation types also call «in-phase» and «anti-phase», «even» and «odd» [4], as well as «oscillator» and «rotator», depending on their «phase portrait» [5]. Therefore, they can be denoted as f_1 and f_2 . Accordingly, the transfer coefficient \mathbf{T} and reflection coefficient \mathbf{R} for two-resonance BSF based on RPC structures have the following form:

$$\mathbf{R} = K_2/(1 + K_2) - K_1/(1 + K_1) \quad (11)$$

$$\mathbf{T} = 1 - K_1/(1 + K_1) - K_2/(1 + K_2) \quad (12)$$

The important property of filters based on RPC structures that \mathbf{T} and \mathbf{R} of bandpass filters are dual to \mathbf{R} and \mathbf{T} of bandstop filters (see (1), (2), (11), (12), respectively). The structures based on the RPC are modeled by bridge equivalent circuits [6]. For bridge equivalent circuits the duality of resistance $[Z]$ and conductivity $[Y]$ matrices is typical [7]. This property is formulated as follows [8]: if we replace only one of the filter arms with a reverse two-pole (parallel and series resonators), the passbands and bandstop will change places, i.e. we obtain an additional filter. As result, a low-pass filter turns into a high-pass filter, and a bandpass filter turns into a bandstop filter and vice versa. Note that this property can be useful in the study of physical processes in metamaterials and structures with Fano resonance. The spectra of electromagnetic waves (from microwave to X-ray wavelengths) at the outputs of studied samples (transmission coefficient \mathbf{T} is absorption spectrum) can to be dual to spectra obtained at the reflection of electromagnetic waves from the same samples (reflection coefficient \mathbf{R}). The duality of the \mathbf{T} and \mathbf{R} characteristics of the structures with Fano resonance is presented in Sect. 4.5 «Fano resonances in non-Hermitian photonic structures» [9].

The conditions for zero reflection bandpass filter (ideally matched), the critical coupling for a bandpass filter (maximally flat) and conditions for infinite damping of the RF are given in [3]:

$$a^2 = K_1 K_2 - 1 \tag{13}$$

$$a = K + 1 \text{ (for } K_1 = K_2 = K) \tag{14}$$

$$a^2 = K_1 K_2 - 1 \tag{15}$$

Transmission coefficients **T** and reflections coefficients **R** showed in Figs. 1 and 2 (amplitude frequency response of BPF and BSF for different ratios between the coupling coefficients *K* and detuning *a*). Transmission coefficient **T** of a bandpass filter fully corresponds to reflection coefficient **R** of a bandstop filter and vice versa, the transmission coefficient **T** of the BSF fully corresponds to the reflection coefficient **R** of a bandpass filter. Another important feature of bridge filters characteristics is the amplitude frequency response of the transmission and reflection coefficients of the filter (blue curve), which lies between amplitude frequency response (green and red dotted curve) of individual resonators. Note that this property is also inherent in the “Fano resonance”. The Fano resonance point is the point, in which its amplitude is zero (or very small) and it lies between the resonant frequencies of individual oscillations (resonators) [9, 10].

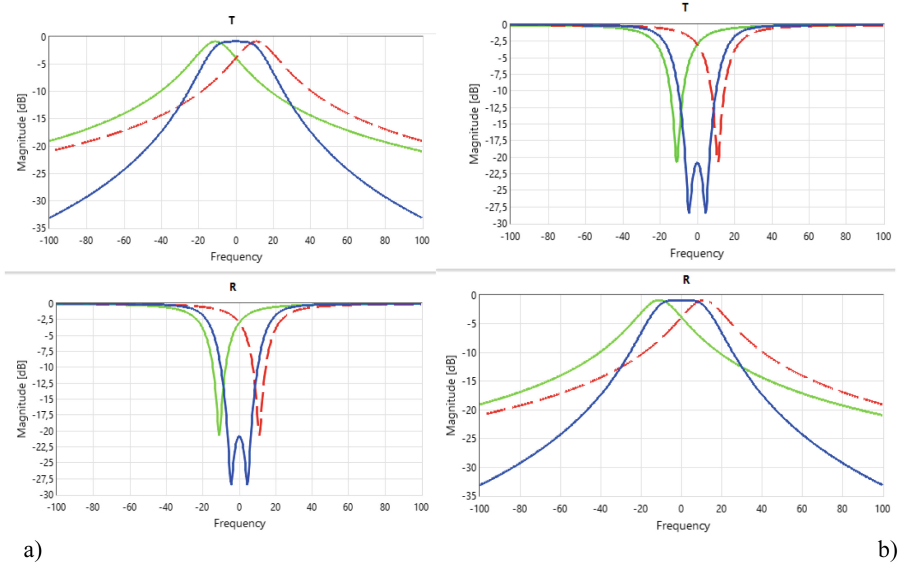


Fig. 1. The transmission coefficient **T** and reflection coefficient **R** (a) bandpass and (b) bandstop filters. $K_1 = K_2 = 10$, $a = 11$ - conditions of the “critical” coupling for a bandpass filter, corresponding to the maximally flat characteristic (Eq. (14)). Here and in the graphs below (unless specifically stated) the green curves correspond to the characteristics of the first resonator with the coupling coefficient $K_1(K_2 = 0)$, the red dash ones correspond to the characteristics of the second resonator with the coupling coefficient $K_2(K_1 = 0)$, the blue curves represent the characteristics of a two-resonators structure (filters or structures with Fano resonance).

Consider the characteristics of the amplitude frequency response of the transmission and reflection coefficients of BPF and BSF when relations (13) and (15) hold (Fig. 2). BPF amplitude frequency response \mathbf{R} (ideally matched), since transmission coefficient \mathbf{T} for BSF is almost zero, $\mathbf{R} < -85$ dB, actually $\mathbf{R} = 0$ (small difference is caused by limited digit capacity of the calculations and the selected axle scales). However, this property of the BPF is not important. Another property of the BSF is much more important. The detuning between the resonant frequencies \mathbf{a} and the coupling coefficients of the resonators with the transmission line that satisfy relation (15) gives us infinite signal rejection even in the presence of losses in resonators. The implementation of infinite attenuation is possible only with the use of bridge circuits and it is impossible for structures modeled as ladder lumped element equivalent circuits. This property of bridge circuits, formulated in terms of the theory of circuits, is due to the possibility of the realization of transmission zeroes on the real axis in the Z -plane [8, 11–16].

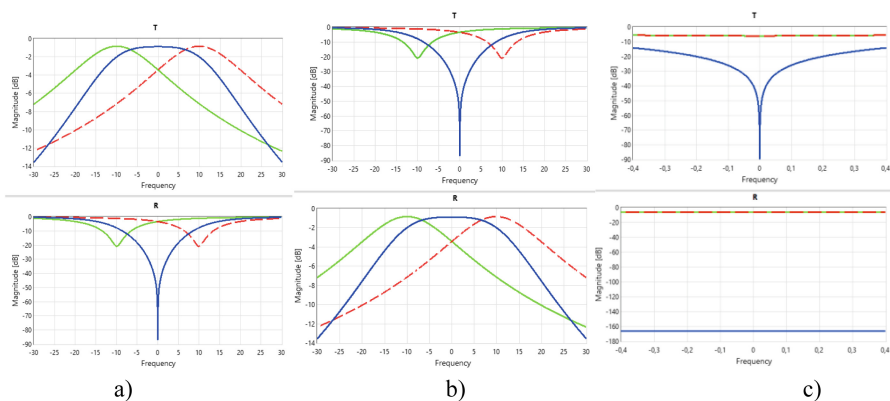


Fig. 2. Amplitude frequency response \mathbf{T} and reflections \mathbf{R} (a) bandpass and (b) bandstop filters. $K_1 = K_2 = 10$, $\mathbf{a} = 9.95$ (the conditions of the “critical” coupling (Eq. (15)) for the bandstop filter; (c) Amplitude frequency response \mathbf{T} and reflections \mathbf{R} of a bandstop filter with extremely narrow locking band $K_1 = K_2 = 1$, $\mathbf{a} = 10^{-8}$.

Note that the BSF is matched at the center frequency with $K_1 = K_2 = 1$ and infinite attenuation (Fig. 2c). The nonzero value $\mathbf{a} = 10^{-8}$ is taken to demonstrate the matching level of such filter, which gives the reflection coefficient $\mathbf{R} \approx -160$ dB. It was noted in [4] if we draw an analogy between coupling coefficients of electric and magnetic type of oscillations K_e and K_m with dielectric ϵ and magnetic μ permeability of a metamaterial, then, under condition $f_e = f_m$ the equality $K_e = K_m = 1$ can be assumed to the equality $\epsilon = \mu = -1$, which in metamaterials corresponds to the parameters of a so-called “superlens” [17]. Obviously, a «superlens» cannot consist of a single metamaterial cell, but it will consist of such resonant cells. The fact that the prototype of the «superlens» is a filtering structure, can be obtained from the characteristics of the spectra of the incident and transmitted signals [18]. In addition, as will be shown

below, the achievement of infinite attenuation with the maximum degeneration of two oscillations of different types (resonant frequencies of oscillations of electric f_e and magnetic f_m types of oscillations coincide) is a special case of Fano resonance.

2 Different Q-Factors of Resonators (Oscillations)

Analysis of the characteristics of BPF and BSF for different Q-factors of the resonators turned out that their amplitude frequency response look like Fano resonance characteristics. Identical Q-factors of resonators in structures with the RPC are the exception rather than the rule, because they operate using different oscillations (electric and magnetic, in-phase and anti-phase even and odd ...). We add the parameter b (expression (10)), which takes into account the ratio between self-quality factors of resonators. Figure 3 presents the amplitude frequency response of two-resonance BSF for the values $b = 4$ and $b = 0.25$, respectively. The detuning values “ a ” are chosen to achieve maximum attenuation at one point (Fano resonance point).

The condition (15) for achieving the maximum attenuation at different Q-factors of the resonators “splits” into two equations:

$$a_{ext1} = \frac{1+b}{2b} \sqrt{(K_1 K_2 - 1)} \quad \text{for } b < 1 \tag{16}$$

$$a_{ext2} = -\frac{1+b}{2b} \sqrt{(K_1 K_2 - 1)} \quad \text{for } b > 1 \tag{17}$$

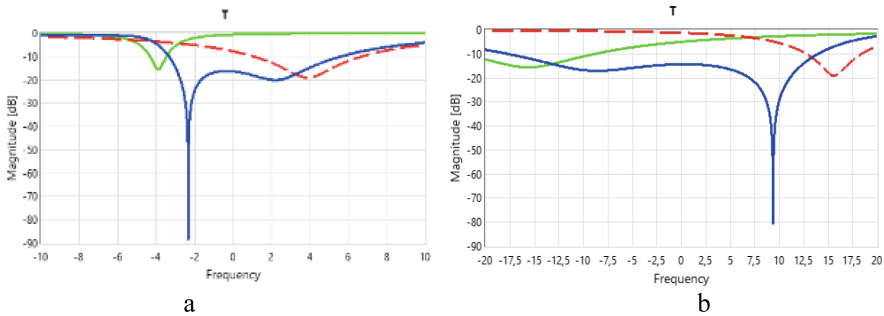


Fig. 3. The coupling coefficients $K_1 = 5, K_2 = 8$ are the same for both graphs. $b = 4, a = 3.903$ in (a) $b = 0.25$ and $a = 15.612$ in (b).

In general, the question of Q-factor of resonators in processes or devices whose characteristics are described by Fano resonance (FR) is very relevant. First, as a rule, this is a pronounced characteristic asymmetry. Secondly, as shown below in Sect. 3, it is one of the reasons for the formation of damping poles in bandpass filters based on RPC structures. Thirdly, different Q-factor is the cause of the asymmetry of electromagnetic elementary cells of metamaterials (Split Ring Resonators - SRR). This

asymmetry was already apparent in one of the first published papers on metamaterials [19], in which the experimental SRR resonant curve in a waveguide was presented. Due to the fact that the cutoff frequency of a waveguide is higher than the half-wave resonance frequency of the outer ring, this first oscillation did not appear in the experiment. The Q-factor of the inner and outer rings cannot be the same, at least because of their different lengths. The outer one is almost twice as long as the inner one. However, the main problem is not in the asymmetry of the resonance characteristic, but it was perceived as a resonance curve of ordinary oscillation. It is now clear that we are talking about the structure described by the Fano resonance. Therefore, the SRR resonance curve is formed by at least two independent oscillations of different types. At a certain (optimal) mutual tuning (or, more precisely, frequency detuning between the resonant frequencies of individual rings) very large attenuation levels are achievable at the Fano resonance point (up to -50 dB in low-Q microwave structures [4], more than 100 dB in optical microresonators [20]). These very high attenuation levels in comparison with ordinary resonant structures are achieved not due to the Q-factor of the resonators, but due to the degeneration of oscillations.

Analysis of the characteristics in Fig. 3 gives several important conclusions:

1. Fano resonance, as well as the SRR attenuation characteristic, is formed by two oscillations of different types, which, at small detuning between their self-resonant frequencies and close values of the Q-factor can appear as one degenerate oscillation.
2. The quality factor of such degenerate oscillation is incorrect to determine. It can be extremely large.
3. The resonance frequency of the total (degenerate) oscillation (blue curve) always lies between resonant frequencies of individual oscillations (“green” and “dash red”). This property of Fano resonance, which we have already mentioned for filters in Sect. 1, is also extended to the resonant characteristics of metamaterial cells. If by analogy with structures (processes) with coupled oscillations we call the initial frequencies of individual oscillations (resonators) in the system with the RPC as “partial” [21], then the total degenerate oscillation of the structure with FR can be considered as a “normal” or “proper” oscillation [22, 23]. Consideration of this problem is Sect. 6.
4. Structures based on the RPC with small (and different) Q-factor of resonators included in a system may well provide infinite attenuation (absorption) at one frequency (Fano resonance point). It is interesting to note that this also applies to acoustics. A system formed by a set of in-phase and anti-phase acoustic resonators was considered in [24] and theoretical possibility of achieving complete oscillations absorption at one frequency was demonstrated. Unfortunately, there are no experimental data of similar research. However, both the theory of Fano resonance and the model based on bridge lumped element equivalent circuits make it possible to describe acoustic oscillatory processes. Therefore the results obtained in [24] seem quite realistic.

3 Bridge Equivalent Circuits for Metamaterial Cells and Structures with Fano Resonance

Metamaterial cells are usually present as a ladder equivalent lumped element circuits model [25–29]. However, as noted in [2], the equivalent lumped element circuits model presented in Fig. 4 are more suitable for this purpose.

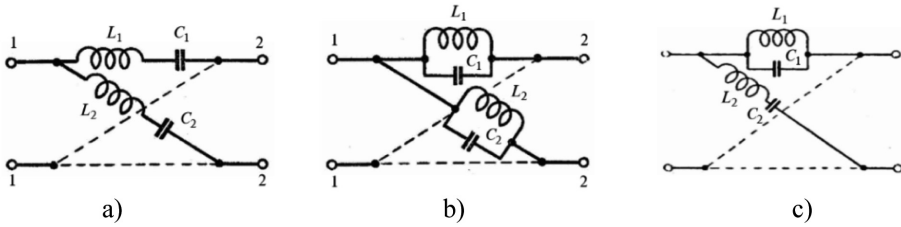


Fig. 4. Bridge equivalent circuits (a)-(b) of the bandpass filters, (c) – of the bandstop filter

Most accurately the electrodynamic cell of metamaterials corresponds to the bridge circuit of Fig. 4c. For structures with SRR, the presence of two degenerate oscillations became not immediately apparent. As for the structures with Fano resonance, the presence of two independent oscillations (wave processes) of different types and independent summation of their energies directly indicate the most suitable version of the circuit model based on lumped elements—the bridge circuit. Recently, there have appeared articles and books in which microwave filters are considered in the terminology of even and odd types of oscillations [30–32]. Bridge equivalent lumped element circuits are not directly considered in them, however, it is obvious that structures with even and odd oscillation types (filters, directional couplers, balanced bridges) can be modeled by bridge circuits [8, 33].

Bridge circuits in circuit theory have been well studied [8, 9, 11–16]. Why are the bridge circuits, or, as they are called non-minimal-phase circuits, chosen by us for modeling Fano resonance? In [11] the following important properties of bridge 4-poles are noted:

The zeros of the function $T(p)$ describing the transmission coefficient of a stable 4-pole can be located both in the left and right half-plane in the Z -plane of the variable p , and on the real axis. It means that if a signal is presented at the input

1. of the 4-pole, at a certain frequency the signal at the output vanishes, i.e. the signal is infinitely rejected. It can not be implemented in lattice lumped element equivalent circuits with minimal losses.
2. With the same number of zeros and poles, the «non-minimum phase» circuit provides a greater in absolute value change in the phase of the transmission coefficient compared to the «minimum phase» circuit. Below in Sect. 4 we will show that the specific properties of the phase-frequency characteristic of the transmission coefficient of the bridge filter lead to extremely large group delay time in this structure. As is known [9], Fano resonance is associated with such physical phenomena as

electromagnetically induced transparency (EIT). Extremely high electromagnetic signal delay is inherent in the links of the bridge filter with a certain choice of its parameters.

3. Any 4-pole can be «minimum phase» with the following property: the signal transmission from input to output can be completely stopped by breaking a single branch of any 4-pole ladder structure. In structures with Fano resonance, there are at least two independent paths of energy distribution; so, ladder schemes are not always suitable for describing their properties.
4. As a rule, nonminimum-phase 4-poles have the structure of bridge (crossed) circuits, in which the signal passes from input to output on two different channels. The same is typical for the Fano resonance.
5. The bridge structure does not automatically guarantee that the circuit belongs to the non-minimal phase class. In each case you should check the presence or absence of zeros of the transfer function in the Z-plane. In microwave structures similar to SRR, there can exist two oscillation degeneration regions depending on the ratio between resonant frequencies of electric f_e and magnetic f_m types of oscillations [34]. It also follows from (15). The detuning a , corresponding to the condition of complete attenuation at the center frequency, can have two different values (the roots of the quadratic equation). The structure is non-minimally phase in one of these areas (where $f_e > f_m$) and it has properties inherent in metamaterials.

If 4-pole is non-minimal-phase then its amplitude frequency response and phase-response are independent of each other. Based on the properties of Hilbert’s transformations we can argue, for example, that if amplitude frequency response of a minimum-phase 4-pole reaches a maximum at some frequency, then the amplitude frequency response in the vicinity of this frequency passes through zero (green and dash red curves in Fig. 5 are individual «branches» of a bridge 4-pole). This does not apply to non-minimum phase circuits (blue curves in Fig. 5). The blue curve does not intersect zero at all.

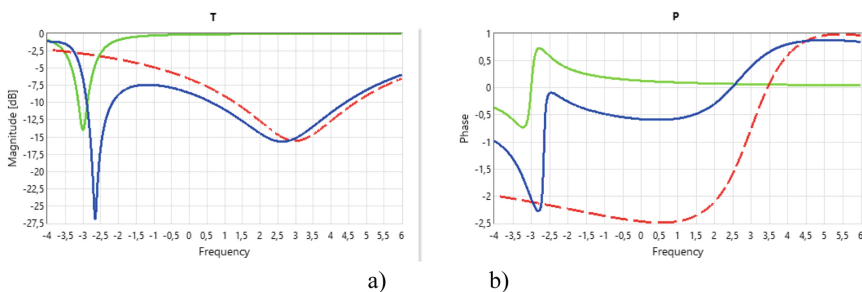


Fig. 5. (a) frequency-response and (b) phase-response characteristics of the transmission coefficient of the bridge bandstop filter (Fig. 4c). $K_1 = 4$, $K_2 = 5$, $a = 3$, $b = 10$.

The parameters of the structure, the characteristics of which are shown in Fig. 5a and b, are chosen arbitrarily in order to demonstrate the absence of an unambiguous relationship between phase response and amplitude frequency response of a

non-minimum phase bridge circuit. The detuning $a = 3$ differs from the detuning defined by Eq. (17), from which it follows that for $K_1 = 4$, $K_2 = 5$, $b = 10$, the damping pole $a_{ext2} = -2.39739$.

Figure 6 demonstrates the characteristics of the transmission coefficient T of the bandstop filter modeled for the values of a , calculated in accordance with (16), (17). Amplitude frequency response in Fig. 6a and b practically coincide in form and values, the phase response is significantly different. In Fig. 6a the phase response passes through zero (jump by 360°) at the point below (in frequency) the minimum amplitude frequency response point. This is clearly visible in relation to the yellow line in the lower graph. In Fig. 6b the phase response passes through zero at the point above (in frequency) the minimum amplitude frequency response point. This is also clearly visible in relation to the yellow line in the lower graph. In this case the phase response changes at the point close to the extremum by 180° . In that model, built on the basis of the equation given in Sect. 1, the attainable attenuation level and related parameters depend on the limited digit capacity of the calculations, and in real physical structures, depend on Q-factors of resonators.

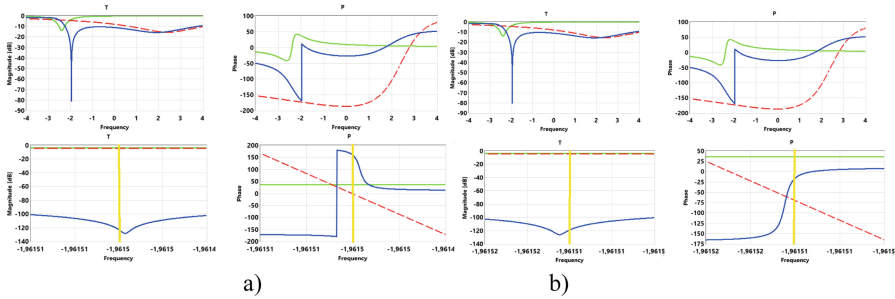


Fig. 6. The amplitude frequency response and phase response of the bridge filter in the vicinity of extreme attenuation, $K_1 = 4$, $K_2 = 5$, $b = 10$; (a) $a_{ext2} = -2.39739$; (b) $a = 2.3974$. The upper line of the table is a small scale (1 unit per division), the lower line is a large scale ($5 \cdot 10^{-6}$ units per division). The maximums of the amplitude frequency response and phase response are shown.

The fundamental property of a non-minimal phase structure is still fulfilled. The transition point of the phase response through zero does not coincide with the point of the phase response extremum. However, it's important to dwell on another characteristic of the behavior of the studied structure in the vicinity of the extremal point considered in [16]:

- (a) Any system that has infinite attenuation at signal frequencies in some finite band, including any ideal filter, is not causal.
- (b) Causality also includes a strong relationship between real and imaginary parts (amplitude and phase) of an amplitude frequency response. In fact, one follows from the other.

It can be seen, the extremely large achievable attenuation and the absence of a strong dependence between amplitude frequency response and phase response suggest that

non-minimally phase bridge structures are not causal. Indeed, as shown the characteristics in Fig. 5, at frequencies close to the extreme (attenuation maximum), any slightest change in structure parameters (at thousandths of a percent, almost at the level of intrinsic thermal noise of the structure), associated, for example, with environmental exposure, can drastically change the phase response of the filter by 180° , that to change the direction of propagation of the signal at the output with respect to the input.

4 Results of Experimental and Theoretical Studies of Bandstop Filter with Resonators in Parallel Channels

Consider abnormally high Q-factors and group delay time as important properties of elementary cells of metamaterials and bridge circuits that model them in more detail. The results of theoretical and experimental studies of microstrip bandstop filter (Fig. 7) [35] are presented below.

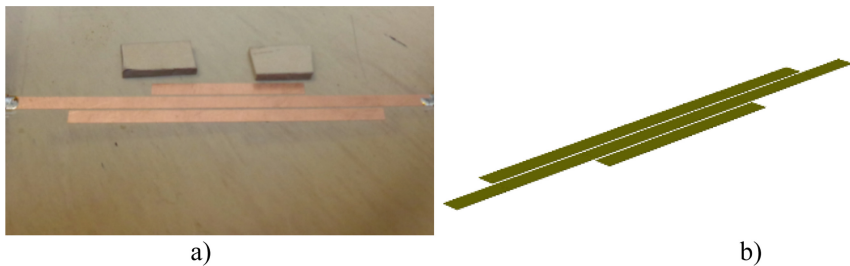


Fig. 7. (a) Two-resonator bandstop filter [35]. Dielectrical permittivity $\epsilon = 10$, substrate thickness $d = 2$ mm, microstrip width is 2 mm, material – copper, long resonator – 42 mm, short resonator – 21 mm, gap between the line and resonators is equal to 0.1 mm. Two pieces of substrate dielectric were shown in Figure. The accurate tuning of the microstrip resonator’s frequencies were carried out using these pieces; (b) filter topology, whose characteristics are calculated using AWR program.

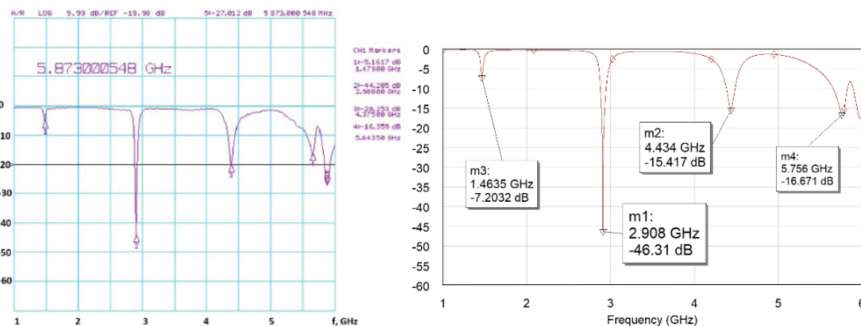


Fig. 8. (a) Experimental and (b) calculated (AWR program) filter (Fig. 7) characteristics Parameters of “marks” in (a) Start 1 GHz, Stop 6 GHz, “mark” 1: -5.1617 dB, 1.47900 GHz; “mark” 2: -44.205 dB, 2.90000 GHz; “marks” 3: -20.253 dB, 4.37500 GHz; “mark” 4: -16.359 dB, 5.6435 GHz; “mark” 5: -27.012 dB, 5.87300 GHz;

Let us compare experimental spectra (Fig. 8a) and spectra calculated using the AWR program in the 1–6 GHz range (Fig. 8b). It is important to note:

- (1) the first harmonic missed in the experiments with metamaterial cells [19] is due to the fact that it has frequency lower than the cut-off frequency of the waveguide;
- (2) the anomalously large attenuation with a relatively narrow frequency band appears where harmonics 1 and 2 of a short and long resonator are close in frequency (near 3 GHz).

It can be seen that the amplitude frequency response of these oscillations, for which one detached harmonic is responsible, is typical for single-resonators structures with a relatively small Q-factor. Differences in the calculated and experimental parameters of higher harmonics (no more than 1–2%) are due to the fact that dielectric samples with low dielectric constant were used when a filter was fine-tuned in the extreme attenuation band. Its resonant frequency changes when they are superimposed on the edges of the resonator. In the calculations (AWR program) the effect of the dielectric samples on the higher harmonics of the oscillations of the resonators was not taken into account.

The Q-factor of the resonator (oscillation) is determined by its amplitude frequency response as the ratio of the center frequency of the passband (stopband) to the band at the –3 dB level, measured from the maximum value of the amplitude frequency response:

$$Q_0 = \frac{f_0}{2\Delta f_{-3\text{ dB}}} \tag{18}$$

The experimental amplitude frequency response of the filter transmission coefficients is shown in Fig. 9: (a) the frequency of “degeneration” of the first harmonic of the half-wave (short) resonator and the second harmonic of the “wave” (long) resonator, and (b) at the frequency of the third harmonic of the “wave” (long) resonator.

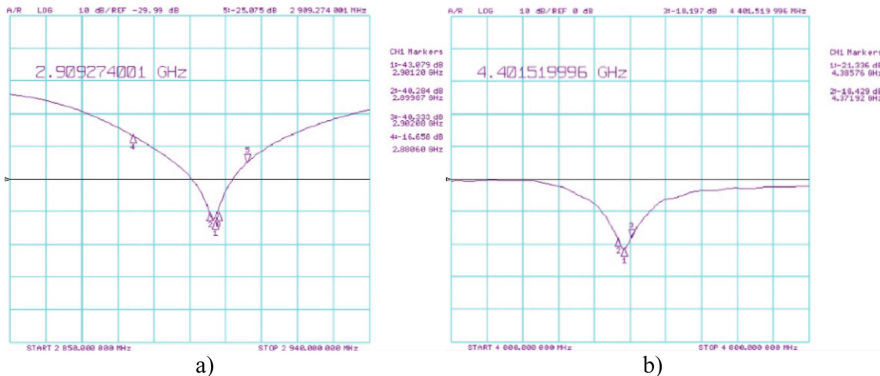


Fig. 9. Experimental amplitude-frequency characteristics of the filter Fig. 6. (a) in the field of degeneration: Start 2.850 GHz, Stop 2.940 GHz; “mark” parameters: “mark” 1: –43.879 dB, 2.90120 GHz; “mark” 2: –40.284 dB, 2.89987 GHz; “mark” 3: –40.333 dB, 2.90208 GHz; (b) in the frequency range of the third harmonic of the “wave” (long) resonator: Start 4.000 GHz, Stop 4.800 GHz; “mark” parameters: “mark” 1: –21.336 dB, 4.38576 GHz; “mark” 2: –18.429 dB, 4.37192 GHz. “mark” 3: –18.197 dB, 4.4015 GHz.

The values of the Q-factor about 150 for ordinary single-resonators oscillations (amplitude frequency response Fig. 8b) were obtained. The band $2\Delta f$ (-3 dB) is determined as double difference between 1 and 2 marks and it is equal about 1300 at 2.902 GHz. Such abnormally high Q-factor inherent in structures like SRR is related to the fact that the parameters of the degenerate oscillation are measured, for which the concept of quality factor, defined by formula (18) is not correct generally. In a single-resonator filter, the quality factor characterizes the ratio of the stored energy to the energy of losses over a period. If we are talking only about unloaded good quality, then the losses are associated only with the losses in the resonator. If we are talking about loaded quality, then the energy given to external communication channels is also taken into account. As can be seen from bridge circuits structure, the stored energy can “run along the contour”, which is formed by parallel and crossed shoulders, and there are two such contours and the directions of movement in them are opposite (“clockwise” and “counterclockwise”). In the case of resonance in parallel arms, the energy can spread over the arms crossed. In the case of an “ideal” bandstop filter, the characteristic of which is shown in Fig. 2c, $T = 0$, $R = 0$, therefore, the energy at the output does not pass and does not reflect from the input of the filter, being stored completely inside the structure (energy is spent only on heat losses in the resonators), therefore it can be “stored” in high-quality optical resonators for a long time (Fig. 10).

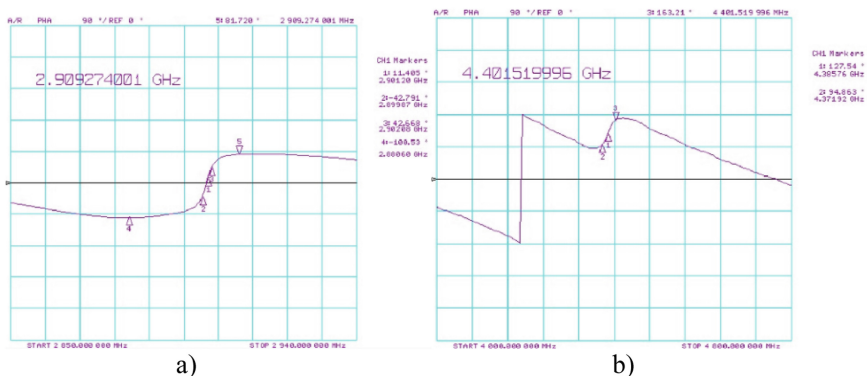


Fig. 10. Experimental phase-frequency characteristics of the filter Fig. 6. (a) in the field of degeneration: Start 2.850 GHz, Stop 2.940 GHz; “mark” parameters: “mark” 1: -11.405° , 2.90120 GHz; “mark” 2: -42.791° , 2.89987 GHz; “mark” 3: 42.668° , 2.90208 GHz; «mark» 4: -180.53 , 2.8805 GHz “mark” 5: 81.720° , 2.90974 GHz (b) in the frequency range of the third harmonic of the “wave” (long) resonator: Start 4.000 GHz, Stop 4.800 GHz; “mark” parameters: “mark” 1: -23.343 dB, 4.38576 GHz; “mark” 2: -18.406 dB, 4.37192 GHz.

The phase difference between marks 4 and 5 is more than 180° and it is due to the additional phase shift between the input and output due to the large length of the resonators. This additional phase shift can be estimated from the phase response outside the resonance band as $1^\circ/\text{MHz}$ (approx 30°). According to Fig. 9a the group delay time GD (Group Delay) of one individual harmonic oscillation is estimated

approximately $2.5\text{--}3^\circ/\text{MHz}$ and GD in the region of 3 GHz is about $50\text{--}100^\circ/\text{MHz}$. As it seen, the studied filter demonstrates the properties, which characterize bridge circuits, an abnormally high Q-factor and extremely high GD.

As already mentioned, calculations using AWR program model experimental characteristics with high confidence, we will use this program to simulate structures with RPC with higher Q-factors than of real microstrip resonators.

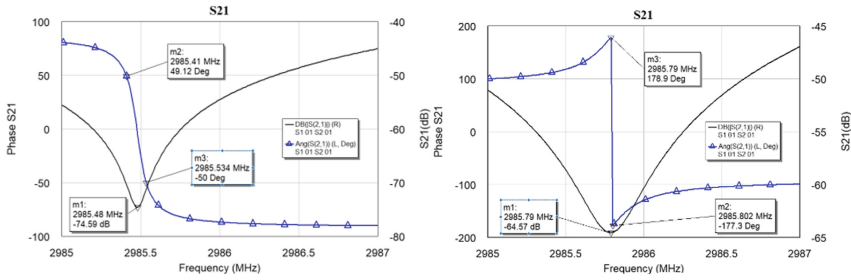


Fig. 11. Black curves is amplitude frequency response, blue curves is phase response. Demonstrated options of “switching” the phase in structures with the RPC upon reaching the maximum attenuation: (a) «jump» on 180° (correspond amplitude frequency response at Fig. 6 (b)), (b) «jump» on 360° (correspond amplitude frequency response at Fig. 6(a)). These characteristics are obtained for the resonator Q-factor is about 10000.

For the Q-factor of the order of 10000 we received amplitude frequency response and phase response that are shown in Fig. 11. The parameters of the structure (resonator lengths) slightly differ from the parameters of the real filter in Fig. 6. The behavior of the phase response near the resonance is similar to Fig. 6. Also, near the extremum point, two completely different scenarios of the behavior of the phase response are observed. We have a transition through zero or jump on 180° . Estimated group delay time is about $800^\circ/\text{MHz}$.

5 Bridge Equivalent Circuits for Microwave Filters with Additional Energy Transmission Channels

As shown above, in bandstop bridge filters modeled by the equivalent circuit in Fig. 4c, the maximum attenuation levels are achieved with detuning between the resonant frequencies, depending both on the level of coupling between the resonators and the transmission lines, and on the relation b (Eq. (10)) and Q-factors of the resonators. Additional energy transmission channels play an important role in the formation of the attenuation poles of the amplitude frequency response of the transmission coefficients of bandpass filters. The formation of attenuation poles in accordance with certain requirements can significantly increase the selective properties of bandpass filters [3, 36–39]. As noted earlier, bandpass filters based on the RPC are modeled by the equivalent circuits in Figs. 4a and b. The same can be said about filters on coupled

resonators, if their normal phase and anti-phase oscillations are considered as separate oscillations corresponding to the branches of the bridge filter. Therefore, consideration of the mechanism of the formation of attenuation poles using equivalent bridge circuits equally applies to both structures based on the RPC, and structures based on coupled resonators.

While in structures with Fano resonance type characteristics we always deal with at least two independent resonant channels for energy transmission, then bridge equivalent circuits for microwave bandpass and bandstop filters also describe structures with additional nonresonant channels for energy transmission [3, 36]. Attenuation poles in bridge filters are formed both due to different reactivities (inductance, capacitance, RL or RC chains) and due to different Q-factors of the resonators in the filter arms.

Consider the first mechanism of the formation of poles on the example of single-resonator BPF. The examples of such structures, formulas for calculating the transfer and attenuation coefficients and their characteristics are presented below. Typical forms of the amplitude frequency response of the transmission coefficient of such filters are presented in Fig. 12:

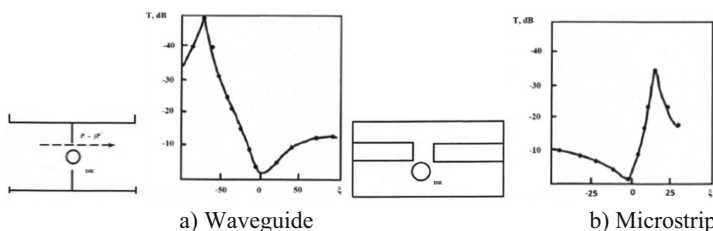


Fig. 12. Bandpass filters on dielectric resonators with non-resonant energy transmission channels: (a) waveguide filter with DR in the diaphragm and its amplitude frequency response; (b) microstrip filter with a gap in the center conductor of the transmission line and its amplitude frequency response.

The transmission coefficient of the structure, taking into account the non-resonant coupling of the input and output transmission lines, is equal **T** bridge circuits. The resonant circuit and the channel of non-resonant energy transmission are located in different branches and their transmission coefficients are summed:

$$T = \frac{K e^{j\varphi}}{1 + K + j\xi} + P' + jP'' e^{j\psi} \tag{19}$$

where **K** is the coupling coefficient of the dielectric resonator (DR) with the input (output) transmission line, determined by the ratio of its unloaded and loaded Q-factors, ξ is the generalized detuning of the resonator (Eq. (9)), the parameter $\varphi = 0$ when the type of oscillations of DR corresponds to an electric dipole and $\varphi = \pi$ —magnetic, $P' + jP''$ is the coefficient of non-resonant energy transmission from the input to the structure output, moreover, with the inductive nature of the junction (diaphragm with a hole in the waveguide, Fig. 12(a)) the parameter $\psi = \pi/2$ jP'' and there is a plus sign in

front of jP'' and for capacitive (gap in the central strip of the microstrip line, Fig. 12(b)) $\psi = -\pi/2$ and there is a minus sign in front of jP'' .

There is a “pole” of attenuation on one of the slopes of the amplitude frequency response of the transmission coefficient T , which location (in the approximation $P' = 0$) corresponds to a generalized detuning

$$\xi_p = e^{j(\varphi - \psi + \frac{\pi}{2})} \left(\frac{K}{2P''} + \sqrt{\left(\frac{K}{2P''}\right)^2 + (1+K)^2} \right) \tag{20}$$

The characteristics of a band-passing structure with two dielectric resonators in a cut-off waveguide, excited on higher types of oscillations with separated resonant frequencies (structures with RPC in the previously accepted terminology) are presented in Fig. 13.

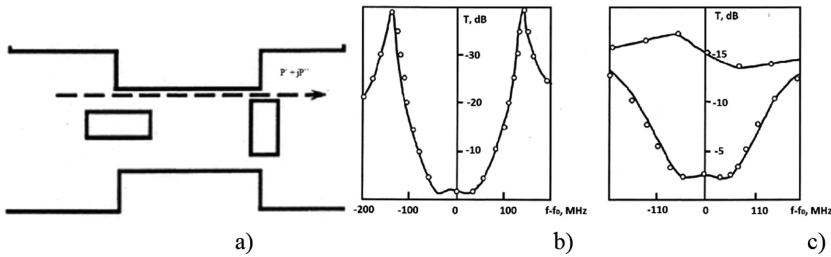


Fig. 13. (a) Two-resonators bandpass filter with DR with different types of oscillations; (b) relative detuning $a > 0$; (c) relative detuning $a < 0$.

The studied structure contains two cylindrical DR, which are located in the segment of the cut-off waveguide and they have a ratio between the cylinder thickness L and diameter D , close to 0.4 ($\frac{L}{D} \sim 0.4$). The resonant frequencies of the $E_{11\delta}$ and $H_{11\delta}$ oscillations are also close, and their relative detuning “ a ” relative to the average frequency $f_0 = \frac{f_c + f_m}{2}$ is determined in accordance with (8). The transmission coefficient of the structure, taking into account the final isolation of the input and output transmission lines takes the form:

$$T = \frac{2jKa}{(1+K+j\xi)^2 + a^2} + jP'' \tag{21}$$

where K is the coupling coefficient with the input and output (defined analogously to formula (5) as the ratio of unloaded and loaded Q -factors taking into account the connection with the input-output of the DR), ξ is the generalized detuning, determined in accordance with formula (8), self-quality factors for oscillations $E_{11\delta}$ and $H_{11\delta}$ may be different. The form of the amplitude frequency response of the transmission coefficient is determined by the detuning sign “ a ”, i.e., the ratio between the natural

resonant frequencies of oscillations. The attenuation poles on two slopes of the amplitude frequency response of the transmission coefficient are physically realizable only when $a > 0$, i.e. when the resonant frequency of oscillations of type $E_{11\delta}$ is less than the resonant frequency of oscillations of type $H_{11\delta}$.

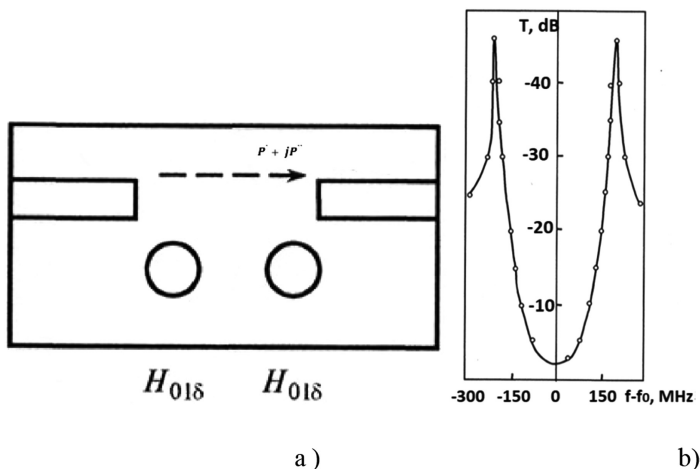


Fig. 14. (a) Band-pass filter with two coupled DR located in the area of the gap of the center conductor of the microstrip transmission line, (b) amplitude frequency response of the filter.

The transmission coefficient of bandpass filters with two interconnected DR, excited on the same type of oscillations $H_{01\delta}$ (Fig. 14a), is described by an expression similar to (21), taking into account the replacement of the generalized detuning “a” by the coupling coefficient between the resonators. Putting $P' = 0$ (large attenuation in the segment of a cut-off waveguide), we get

$$T = \frac{2jK K_c}{(1 + K + j\xi)^2 + K_c^2} \pm jP'' \tag{22}$$

where K is the coupling coefficient of resonators with the nearest transmission line, ξ is the generalized detuning, K_c is the coupling coefficient between resonators. The attenuation poles on the slopes of the amplitude frequency response of the transmission coefficient are determined by generalized detuning

$$\xi_p = \pm \sqrt{\frac{2KK_c}{P''}} \tag{23}$$

Consider the mechanism of the formation of poles at different Q-factors of the resonators. The possibility of forming attenuation poles on a particular slope of the amplitude frequency response of the transfer coefficient depending on the parameter m , equal to the square root of the ratio of capacitances parallel to L_1 and crossed L_2 shoulders (Fig. 4b) is discussed in [8]. Since the Q-factors of the circuits are related to their reactivity, it should be expected that the position of the attenuation poles depends on the Q-factors. The mechanism of this phenomenon is quite simple (Fig. 15).

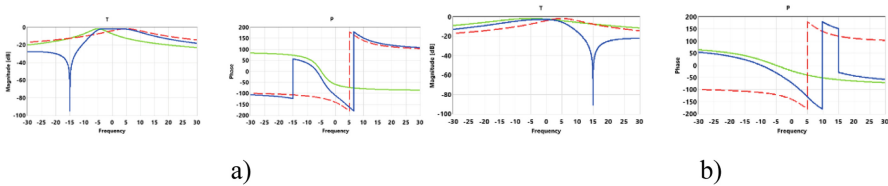


Fig. 15. The graphs are plotted for the values $K_1 = K_2 = a = 5$, (a) $b = 2$ and (b) $b = 0.5$

The pole appears where the transmission coefficients of the “partial” oscillations are equal in amplitude, but opposite in phase (the intersection point of the “green” and “dash red” curves). We have the phase and anti-phase oscillations. The coordinates of the poles are determined by formulas (16) and (17). Naturally, at the pole point the phase characteristic of the filter as a whole has a jump of 180° (changes sign).

6 Normal and Partial Oscillations in Structures with Resonators in Parallel Channel and Structures with Fano Resonance

At the beginning of monumental work “Lectures on the Theory of Oscillations” [40], Mandel’shtam noted, quoting Planck: “correct classification is already a high type of knowledge”. To understand the generality of the processes described by expressions (11)–(12) obtained for bridge equivalent circuits and that described by Fano resonance, we turn to terms common to all resonant processes. Consider the so-called partial and normal oscillations and the relationship between them in structures with coupled oscillators and in structures with the RPC. For coupled oscillations we have the following [21]:

1. Any conservative linear system with n degrees of freedom can be represented as a set of n noninteracting oscillators. This means that a linear conservative system with constant parameters is completely characterized by the spectrum of normal frequencies. This property also can be formulated as follows: normal modes independently of one another contribute to the energy of the oscillatory system [41].
2. A partial system corresponding to a given coordinate is a system obtained from the initial “fixing” of all other coordinates. The equations for finding the partial frequencies can be obtained from the general equations of the system of coupled

oscillations, removing from them the terms expressing the connection between the oscillations.

3. Partial frequencies always lie between normal frequencies. This property of coupled oscillations follows from the fundamental theorem of quantum physics [42] (see also [43]). We are forced to refer here to the Russian version of the article in Wikipedia, since it is precisely in it that parallels are made between quantum-mechanical processes and waves. In particular, “the quasi-intersection of energy levels” is interpreted as “partial frequencies always lie between normal ones”.
4. There is another very important for us property of coupled oscillations [42]. The frequency of in-phase oscillations (phase mode) is always lower than the frequency of anti-phase oscillations (anti-phase mode).

Now turn to structures with resonators in parallel channels that are not coupled. The initial frequencies of each of the resonators at the same time are partial, if we follow the classical terminology adopted for coupled resonators (green and red curves in Figs. 1 and 2). We note that the partial frequencies in the structures with the RPC correspond to the normal frequencies in the structures with split ring resonators (SRR). As noted earlier, the generalized detuning “ a ” between the frequencies of the original oscillations of the resonators corresponds to the coupling coefficient between the resonators in a SRR. Indeed, the more K_c in the SRR the greater the detuning between normal oscillations. The equivalence of K_c and “ a ” also follows from the expressions for the transfer coefficients of the BPF filters on the RPC and SRR (formulas (21) and (22)). It is noteworthy that the Fano resonance frequency lies between the frequencies of the initial oscillations of the system [9, 10]. Thus, it is possible to formulate an essential property of oscillatory systems based on the RPC, structures with the properties of Fano resonance and metamaterial cells—in such structures, normal oscillations lie inside the partial ones.

The second significant difference between oscillatory systems based on the RPC and structures with Fano resonance is that the frequencies of phase oscillations can lie above the frequencies of anti-phase oscillations $f_{even} > f_{odd}$, because they are set when designing the structure and can be rearranged independently [4].

Another interesting property of the bridge circuit is related to the fact that it can describe processes in which independent oscillations are represented as an “oscillator” and a “rotator” [4]. Parametric oscillations are connected with these processes from a children’s toy, a swing and a pendulum with a variable length [41] to the use of the second harmonic (even and odd oscillations) in parametric amplifiers [44].

Considering physical phenomena associated with Fano resonance (electromagnetically induced transparency, interference phenomena, etc.), terms such as bright mode in the Fano-resonant system are used [9]. It is important to search for a correspondence between partial and normal frequencies, even and odd modes from the theory of circuits, “oscillators” and “rotators” of physical processes [5] and “modes” in systems with Fano resonance. Understanding their relationship (it is possible that the same oscillations or modes are differently called in different physical phenomena) is important primarily because the “Fano resonance” is positioned as invariant with respect to the types of oscillations and waves. The same can be said with reference to

analytical expressions for **T** and **R** bridge filters applied to oscillatory systems of different physical nature.

The resonators in a SRR. Indeed, the more K_c in the SRR the greater the detuning between normal oscillations. The equivalence of K_c and “**a**” also follows from the expressions for the transfer coefficients of the BPF filters on the RPC and SRR (formulas (21) and (22)). It is noteworthy that the Fano resonance frequency lies between the frequencies of the initial oscillations of the system [9, 10]. Thus, it is possible to formulate an essential property of oscillatory systems based on the RPC, structures with the properties of Fano resonance and metamaterial cells—in such structures, normal oscillations lie inside the partial ones.

The second significant difference between oscillatory systems based on the RPC and structures with Fano resonance is that the frequencies of phase oscillations can lie above the frequencies of anti-phase oscillations $f_{even} > f_{odd}$, because they are set when designing the structure and can be rearranged independently [4].

Another interesting property of the bridge circuit is related to the fact that it can describe processes in which independent oscillations are represented as an “oscillator” and a “rotator” [4]. Parametric oscillations are connected with these processes from a children’s toy, a swing and a pendulum with a variable length [41] to the use of the second harmonic (even and odd oscillations) in parametric amplifiers [44].

Considering physical phenomena associated with Fano resonance (electromagnetically induced transparency, interference phenomena, etc.), terms such as bright mode in the Fano-resonant system are used [9]. It is important to search for a correspondence between partial and normal frequencies, even and odd modes from the theory of circuits, “oscillators” and “rotators” of physical processes [5] and “modes” in systems with Fano resonance. Understanding their relationship (it is possible that the same oscillations or modes are differently called in different physical phenomena) is important primarily because the “Fano resonance” is positioned as invariant with respect to the types of oscillations and waves. The same can be said with reference to analytical expressions for **T** and **R** bridge filters applied to oscillatory systems of different physical nature.

7 Conclusions

During last three decades a model of uncoupled resonators of different types located in parallel channels is used for the analysis and synthesis of microwave filters. This model is represented by analytical equations and by an equivalent circuit in the form of a bridge circuits. The introduction of a new parameter associated with different Q-factors of resonators (oscillations) made it possible to describe a specific mechanism for the formation of damping poles in microwave bandpass filters and to implement characteristics similar to Fano resonance. The experimental data presented confirm the possibility of implementing structures with abnormally high group delay times and attenuation levels (Fano point). It is shown that simple analytical expressions and equivalent bridge filter circuits make it possible to conveniently simulate and analyze complex physical processes, which include electromagnetically induced transparency

and other physical processes in acoustics, optics and mechanics. The analogy between bridge equivalent circuits and real physical processes in the field of Fano resonance deserve special attention. For example, it is the non causality and the absence of one-to-one correspondence between the amplitude frequency response and phase response characteristics of oscillatory systems and phase jumps of 360° (in fact it is the crossing of the zero level) and 180° (change the sign of the phase response) degrees. A change in the phase of a signal by 180° can be viewed as a change in the direction of the phase velocity V_p . The opposite direction of the group velocity V_g and phase velocity V_p is a characteristic feature of “metamateriality”.

References

1. Ilchenko, M.E., Zhivkov, A.P.: UHF devices based on several dielectric-cavity mode types. *Radioelectron. Commun. Syst.* **32**(5), 56–59 (1989)
2. Ilchenko, M.E., Zhivkov, A.P., Orlov, A.T.: Filters based on resonators with modes similar in frequency as cells of metamaterials. *Naukovi Visti NTUU “KPI”* **104**(1), 7–14 (2016). (in Russian)
3. Ilchenko, M.E., Zhivkov, A.P.: Microwave filters based on the structures with resonators in parallel channels as metamaterial cells. *KPI Sci. News* **6**, 7–21 (2018)
4. Ilchenko, M.E., Zhivkov, A.P.: Properties of metamaterials in oscillation terminology. In: 2017 International Conference on Information and Telecommunication Technologies and Radio Electronics, UkrMiCo 2016, Ukraine, 11–15 September, pp. 170–173 (2017)
5. Sommerfeld, A.: *Atomic Structure and Spectral Lines*, vol. 1. State Publishing House of Technical Literature, Moscow (1956)
6. Ilchenko, M.E., Zhivkov, A.P.: Two-channel microwave bandpass filter. *Elektronnaya Tekhnika Elektronika SVC* (9), 12–14 (1989)
7. Guillemin, E.A.: *Synthesis of Passive Networks: Theory and Methods Appropriate to the Realization and Approximation Problems*. Wiley, Hoboken (1957)
8. Bosyy, N.D.: *Electric Filters*. Gos. izd-vo tekhn. lit-ry USSR, Kyiv (1959)
9. Kamenetskii, E., Sadreev, A., Miroshnichenko, A.: *Fano Resonances in Optics and Microwaves*. Springer, Heidelberg (2018)
10. Tribelsky, M.I.: *Fano resonances in quantum and classical mechanics*. MIREA (2012)
11. Baskakov, S.I.: *Radio circuits and signals*. M. Higher School (1998)
12. Temesh, T., Mitr, S.: *The modern theory of filters and their design*, p. 560 (1977)
13. Bessonov, L.A.: *Theoretical Foundations of Electrical Engineering*. Electrical Circuits, 10th edn. Gardariki, Moscow (2002)
14. Karni, S.: *Network Theory: Analysis and Synthesis*. Allyn and Bacon, Boston (1966)
15. Ulakhovich, D.A.: *Fundamentals of the theory of linear electrical circuits*, Petersburg, p. 816 (2009)
16. Siebert, W.M.: *Circuits, Signals, and Systems*. The MIT Press, Cambridge (1985)
17. Pendry, J.B.: Negative refraction makes a perfect lens. *Phys. Rev. Lett.* **85**(18), 3966–3969 (2000)
18. Fang, N., et al.: Sub-diffraction-limited optical imaging with a silver superlens. *Science* **308**, 534–537 (2005)
19. Smith, D.R., et al.: Composite medium with simultaneously negative permeability and permittivity. *Phys. Rev. Lett.* **84**(18), 4184–4187 (2000)
20. Gorodetsky, M.L.: Giant-quality optical microresonators, p. 415. Fizmatlit (2011)

21. Rabinovich, M.I., Trubetsov, D.I.: Introduction to the theory of oscillations and waves, p. 560 (2000)
22. Crawford Jr., F.S.: Waves (Berkeley Physics Course, Volume 3). Mcgraw Hill College, New York City (1968)
23. Landau, L.D., Lifshitz, E.M.: Mechanics: Volume 1 (Course of Theoretical Physics S). Butterworth-Heinemann, Oxford (1976)
24. Lazarev, L.A.: Panels with low-Q resonators that have a theoretically infinite soundproofing ability at one frequency. *Acoust. J.* **61**(4), 522–528 (2015)
25. Baena, J.D., et al.: Equivalent-circuit models for split-ring resonators and complementary split-ring resonators coupled to planar transmission lines. *IEEE Trans. Microw. Theory Tech.* **53**(4), 1451–1461 (2005)
26. Gay-Balmaz, P., Martin, O.J.F.: Electromagnetic resonances in individual and coupled split-ring resonators. *J. Appl. Phys.* **92**(5), 2929–2936 (2002)
27. Baena, D., Bonache, J.: Equivalent-circuit models for split-ring resonators and complementary split-ring resonators coupled to planar transmission lines. *IEEE Trans. Microw. Theory Tech.* **53**, 1451–1461 (2005)
28. Capolino, F.: *Metamaterials Handbook*. Taylor & Francis Group/CRC Press, Boca Raton/Milton Park (2009). Vols. 1, 2
29. Garrido, L.: Classical analog of electromagnetically induced transparency. *Am. J. Phys.* **70**, 37 (2002)
30. Luo, S.: A dual-band ring-resonator bandpass filter based on two pairs of degenerate modes *IEEE Trans. Microw. Theory Tech.* **58**, 3427–3432 (2010)
31. Lin, L., Wu, B.: A novel quad-mode resonator and its application to dual-band bandpass filters. *Prog. Electromagn. Res.* **43**, 95–104 (2013)
32. Ferran, M., Zhu, L.: *Balanced Microwave Filters*, p. 688. Wiley-IEEE Press, Hoboken (2018)
33. Helszajn, J.: *Passive and Active Microwave Circuit*, p. 274. Wiley, Hoboken (1978)
34. Ilchenko, M.E., Zhivkov, A.P.: Areas of degeneration oscillations in metamaterial cells. In: 2017 International Conference on Information and Telecommunication Technologies and Radio Electronics, UkrMiCo, pp. 1–4 (2017)
35. USSR Inventor's Certificate 1529321
36. Ilchenko, M.: *Selected Works*. Naukova Dumka, p. 72 (2011)
37. Zakharov, A., Ilchenko, M.E.: Planar three-resonator bandpass filters with cross coupling. *J. Commun. Technol. Electron.* **62**, 185–193 (2017)
38. Zakharov, A., Ilchenko, M.E.: Thin bandpass filters containing sections of symmetric strip transmission line. *J. Commun. Technol. Electron.* **58**, 728–736 (2013)
39. Zakharov, A., Ilchenko, M.: Two types of trisection bandpass filters with mixed cross-coupling. *IEEE Microw. Wirel. Compon. Lett.* **28**, 585–587 (2018)
40. Mandelstam, L.I.: *Lectures on the theory of oscillations*. Science (1972)
41. Pippard, A.B.: *The Physics of Vibration*, p. 656. Cambridge University Press, Cambridge (1989)
42. von Neumann, J., Wigner, E.: Uber das Verhalten von Eigenwerten bei adiabatischen Prozessen. *Phys. Z.* **30**, 467–470 (1929)
43. Wikipedia. https://ru.wikipedia.org/wiki/%D0%9A%D0%B2%D0%B0%D0%B7%D0%B8%D0%BF%D0%B5%D1%80%D0%B5%D1%81%D0%B5%D1%87%D0%B5%D0%BD%D0%B8%D0%B5_%D1%8D%D0%BD%D0%B5%D1%80%D0%B3%D0%B5%D1%82%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B8%D1%85_%D1%83%D1%80%D0%BE%D0%B2%D0%BD%D0%B5%D0%B9. Accessed 19 Jan 2019
44. Louisell, H.: *Coupled mode and parametric electronics* (1960)

Author Index

A

Avdieienko, Hlib, 239

B

Baranov, Alexander, 3

Biriukov, N. L., 205

G

Globa, L., 76

Guetter, Dietbert, 14

I

Ilchenko, M., 121, 278

Ivanov, S., 262

K

Kaidenko, M. M., 227

Kapshtyk, S., 146

Kaydenko, M., 121

Kononova, I., 101

Kravchuk, S., 121

L

Luntovskyy, Andriy, 14

M

Mogylevych, D., 101

Moshynska, A. V., 167

N

Narytnyk, T., 146

Nesterenko, M., 76

O

Osypchuk, S. A., 167

Osypchuk, S., 186

P

Podlipaiev, V., 34

Prychodniuk, V., 34

R

Romanov, O., 76

Roskoshnyi, D. V., 227

S

Shmihel, B. O., 167

Skulysh, M., 76

Solyanikova, V., 186

Stryzhak, O., 34

T

Triska, N. R., 205

U

Uryvsky, L. A., 167

Uryvsky, L., 186

Y

Yakornov, Yevhenii, 239

Z

Zhivkov, A., 278