NATIONAL ACADEMY OF SCIENCES OF UKRAINE

INSTITUTE FOR INFORMATION RECORDING

A.G. Dodonov, D.V. Lande

VITALITY OF INFORMATION SYSTEMS

KYIV NAUKOVA DUMKA 2011

UDC 004.5
BBC 22.18, 32.81, 60.54
J95

A.G. Dodonov, D.V. Lande. Vitality of information systems. – Kyiv: Nauk Dumka, 2011. – 256 p. (English translation with google translate)

A special place among the tasks that have gained relevance at the present time is occupied by tasks related to ensuring the survivability of information systems.

Within the framework of this work, the information system is considered as a set of meaningfully related elements of the information space. The issues of survivability, modeling of destructive influences, restoration of information network structures are considered, tasks relevant for the functioning of information systems are formulated, and mathematical models are described.

For a wide range of specialists in the field of information technology and information security, senior students, graduate students.

Recommended for publication by the Academic Council of the Institute for Information Recording of the National Academy of Sciences of Ukraine (Protocol No. 16 dated December 13 , 2010)

Reviewers:
A.Ye. Litvinenko – doctor of technical sciences, professor.
V.V. Mohor – Doctor of Technical Sciences, Professor.

Scientific Publishing Department of Physics and Mathematics
and technical literature

Editor *V.V. Verotskaya*

ISBN 978 — 966 — 00 — 1087 — 1

TABLE OF CONTENTS

FOREWORD

Complex systems have various properties, but among them there are those that cannot be neglected when studying their functioning, predicting development, analyzing their interaction with the external environment. These properties include, in particular, survivability. The survivability of a system is understood as its ability to adapt to new unforeseen conditions of functioning, to resist undesirable influences while implementing the main function. Survivability in the traditional sense is a fundamental property of complex systems. Biological, social and many other systems initially have the property of survivability, which allows them to maintain integrity, perform their functions and develop despite degradation, regardless of the presence of adverse (destructive) influences from the external environment [1]. Methods and means of ensuring survivability are used in the creation of complex artificial objects, in particular information systems.

Living systems are able to maintain the continuous performance of their main functions, temporarily or permanently refusing to perform less important functions, change their structure and behavior, find and perform new functions necessary to successfully resist adverse effects, adapting to the conditions of their functioning [2]. The survivability mechanisms included in such systems are an integral part of them, and the evolution of systems determines the evolution of their survivability mechanisms. Thus, the development of systems is a factor in the development of mechanisms to ensure survivability.

Within the framework of this book, an information system is considered as a set of elements of the information space that are meaningfully linked into a network. Accordingly, the properties of this network related to survivability are analyzed, including vulnerability as a characteristic of the violation or preservation of connectivity during a destructive impact on individual components or connections.

Of course, an information system cannot be considered either a biological or a technical system, although individual elements of these systems are necessary for their existence. Most likely, information systems can be attributed to communication systems, the formation of which is significantly influenced by the so — called “human factor”, which is most difficult to formalize.

The main goals of the functioning of information systems include informing (for example, informing about the most important aspects of a particular process or event). At the same time, individual information systems, interacting with each other, are elements of the information space.

It is known that systems are generally divided into targeted and non — targeted. Information systems can belong to both the first and second class. The survivability of artificially generated information systems is crucial, for example, during advertising campaigns and other information operations [3].

This book discusses the theoretical issues of the survivability of information systems, the modeling of destructive effects on them, the restoration of information network structures after undesirable effects. At the same time, elements of the theories of reliability, nonlinear dynamics, percolation, complex networks, phase transitions, etc. are considered, and tasks that are relevant for the functioning of information systems are formulated.

The book is aimed at a fairly wide range of readers: specialists in the field of information technology and information security, students, graduate students; I would like to believe that it will also be useful to analysts who, when solving problems in various fields, want to take into account the features of the modern network information space. We hope that the book will be useful in the preparation of training courses on theoretical and practical issues of survivability of information systems.

The authors express their sincere gratitude to their colleagues and co — authors in the works, fragments of which are summarized in this book: A.A. Snarsky, M.G. Kuznetsova, E.S. Gorbachik, S.M. Braichevsky, V.N. Furashev, V.G. Putyatin, as well as reviewers A.yE. Litvinenko and V.V. Mohor for constructive discussions and comments.

Alexander Dodonov,
Dmitry Lande
December 14th 2010.

INTRODUCTION

In the traditional sense, an information system includes infrastructure elements such as computers, communication lines, etc., as well as an information component (libraries, media, television, radio, billboards, and much more), i.e. represents some meaningful integral subset of the information space.

In the narrow framework of this work, in contrast to the above traditional approach, an information system is considered as some semantic entity, a set of information blocks, an information cluster, the signs of a grouping of elements of which are not always known in advance. At the same time, information systems arise and develop in the information space.

Under the information space we will understand the totality of elements that form some information unity. In this case, the elements of the information space can be a wide variety of information resources: documents that summarize the most diverse types of information – files, emails, web pages, regardless of the types and formats of their presentation. These information resources, in turn, can be grouped into information systems according to semantic features.

The above definition of the information space is qualitative. Of course, the term "space" in this case does not coincide with the concept of "space" in mathematics or physics. Within the framework of this work, space is understood as a kind of common container for the information objects under consideration. As examples of successful theoretical models of the information space, one can cite the “vector — spatial” model of G. Salton or the Barton–Kebler information aging model. The web space model was built at the end of the 20th century by A. Breder [4].

Information systems in our understanding can be interpreted as content systems (English *content* – content), which, as is known [5], fully satisfy the general definition of systems. As special cases of information systems can be considered, for example, thematic collections of documents, websites dedicated to

related to some issues, or information clusters (plots) – arrays of information messages published on various websites dedicated to one topic or one event [5]. With this approach, information systems actually cover a set of elementary content units, which, in the narrow framework of this work, we will call “documents” without distinguishing between the concepts of “document”, “message”, “publication”, etc.

The “habitat” of information systems today is quite representatively represented by the web space, which, however, should not limit the range of issues to consideration of only this network.

This book is devoted to such a property of information systems as survivability, i.e. the possibility of realizing the purpose of functioning in cases of adverse effects. In case of adverse impacts on the information system, some part of the data may turn out to be unattainable, and there may be no possibility of restoring access to this data. The mechanisms for ensuring survivability in this case can be very diverse. For example, a new information process can be generated that realizes the purpose of the functioning of the information system. Such an opportunity will characterize the information system as tenacious. The technologies of such dynamic functional rearrangements of systems are based on the mechanisms of adaptation, compensation, recognition, reconstruction, reconfiguration and reorganization. These mechanisms are used to support the availability, integrity and confidentiality of information at the level necessary to achieve the goal of functioning with a given quality, i.e. survivability is the necessary quality of systems focused on analytical technologies, the most important condition for achieving the goals of which is to provide users with accurate and most complete information at any time and where this information is needed.

Obviously, the survivability of an information system can, on the one hand, be considered as its objective property, which depends on the subject, audience, time, and on the other hand, as a characteristic that they want to give it in case of artificial formation, for example, during information operations.

The property of survivability allows a complex system to remain intact in extreme conditions for it, to adapt to them, changing behavior, structure, and often the purpose of functioning [6]. Vitality is difficult to notice under normal operating conditions . This property manifests itself in relief only in cases of violations in the structure of the system, failure of its components, functional violations, and purposeful destructive influences. Depending on the class of systems, their complexity, degree of organization, as well as on the chosen level of analysis, the survivability property can be assessed as stability, reliability, adaptability, fault tolerance.

The stability of a system is its ability to return to its initial state after the end of the influence that brought the system out of this state, it is the active preservation of certain characteristics by the system, regardless of whether they play any role in the overall system.

Reliability is a complex property of the system, which consists in its ability to perform (under certain operating conditions) the specified functions, while maintaining its main characteristics within certain limits. The most common reliability metrics, which are usually probabilistic in nature, are probabilities of failure — free operation, mean time between failures, availability, etc.

Fault tolerance is the property of a system to remain operational in the event of failure of one or more components.

Adaptability is the property of a system to adapt to changing conditions of the internal and external environment by using various adaptation mechanisms.

Certain properties characterize the system in terms of the possibility of maintaining its functionality in the event of changes in the internal and external environment. If reliability characterizes the functionality of a system under well — defined conditions, stability – under conditions of minor deviations, fault tolerance – in the presence and accumulation of various failures, then survivability – in any, including unforeseen, conditions, in particular, under deliberate destructive impacts.

This book consists of six main sections, the first of which provides a definition of the information system, discusses its properties associated with goals and functions. The first section is also devoted to information space modeling and information systems, one of the most popular approaches is multi — agent modeling. The main models within the framework of this approach are considered.

The second section is devoted to the analysis of survivability, it considers the properties of functional and structural survivability, provides various approaches to assessing survivability, in particular, game — theoretic, entropy, based on logical — probabilistic models, etc.

Modern information systems are network structures of meaningfully related components – elements of the information space. The third section is devoted to the modern direction – the theory of complex networks, within which such important network parameters as clustering, mediation, vulnerability are evaluated from the point of view of survivability, models of "small worlds", properties of scale — free networks, networks with weak links are considered.

The issues of survivability of many networks are associated with such a property as “leakage”, which is studied in the framework of the theory of percolation. The percolation threshold in many cases can be interpreted as the “survivability threshold” of an information system. The fourth section is devoted to the critical level of survivability of information network structures, phase transitions, modeling attacks on networks.

The fifth section deals with a wide range of issues related to information flows from the point of view of the theory of survivability. A formal definition is given, various models of information flows are considered, the results of nonlinear modeling are discussed in detail, as well as the use of modern methods for analyzing time series corresponding to information flows.

The sixth section is devoted to the issues of survivability of information operations, considered as social procedures, the instrument of which is information systems. The stages of information operations, the features of their planning and modeling, as well as methods of monitoring and analysis are described in detail. It provides a large number of practical examples.

I would like to emphasize that the currently observed process in the field of intellectualization of automated systems, the transition from simple data processing to decision support processes required new approaches. In addition, the initial paradigm of information systems, formed several decades ago, no longer corresponds to the real situation – the volume and dynamics of information flows, network topology. It is necessary to search for new principles within the framework of which it would be possible to design qualitatively new systems for processing large and dynamic arrays of information – information systems. A special place among the tasks that have gained relevance in this case is occupied by tasks related to ensuring the survivability of information systems.

The purpose of this book is to systematically present the state of existing theoretical and technological possibilities, present possible development prospects, and give impetus to new ideas in the field of information systems survivability — meaningfully related clusters in the information space.

1. INFORMATION SYSTEMS

1.1. Properties of information systems

In order to compare the properties of information systems with the properties of other systems, let us turn to the classical definition, according to which a system is a set of objects and connections between them, isolated from the environment for a certain time and with a specific purpose. The system in a broad sense is considered as a dynamically changing set of strongly connected objects, which has the properties of organization, connectivity, integrity and segmentation.

Accordingly, the information system in our understanding can be interpreted as a content system, a set of documents connected by mutual contextual links, hyperlinks, citations, common vocabulary, factography, etc., isolated from the environment (information space) for a certain time (relevance time) with for a specific purpose or for a specific reason. Those. Indeed, an information system is a collection of strongly connected objects that has the properties of organization, coherence, integrity (determined by the subject or event) and segmentation (into separate documents, their fragments).

Consider the properties of systems associated with goals and functions.

1. Synergy – the unidirectional action of the components enhances the efficiency of the system. In the case of information systems, in particular information stories, the focus of individual documents enhances the information function of the entire information system.

2. The priority of the interests of the system over the interests of its components (the general thematic trend is determined by the entire information system, and not individual documents as components).

3. Emergence – the goals (information functions) of the components (individual documents) of the system do not always coincide with the goals (functions) of the entire information system.

4. Multiplicity – both positive and negative effects of the functioning of the system components have the property of multiplication, not addition (analogues – the amount of information in documents, information entropy).

5. Purposefulness of information systems in the case of their artificial formation (at the same time, there are non — purposeful systems, including information plots).

6. Alternative ways of functioning and development. The most important documents can be relevant for a long time or individual documents on the same topic, generated in large numbers, but having a short period of relevance.

Structure — related properties of information systems are as follows:

1. Integrity – the primacy of the whole information system in relation to its individual elements (in the general case – documents).

2. Nonadditivity – the fundamental irreducibility of the properties of an information system to the sum of the properties of its constituent elements (of course, if the information system consists of more than one document).

3. Structurality – it is possible to decompose an information system into components (documents), establish links between them.

4. Hierarchy – system components (information messages, documents, perhaps, except for elementary single — aspect ones) can also be considered as subsystems of an information system.

Information systems, like individual documents, are parts of the information space, and, accordingly, they have the following properties associated with the external environment:

1. Communication – the existence of a complex system of communications between an information system and an information space, in particular, individual documents from an information system can be associated not only with other documents from the same information system, but also with other parts of the information space.

2. Interaction and interdependence of the information system and information space.

3. Adaptability – the desire for a state of stable equilibrium, which involves the adaptation of the parameters of the information system at certain stages of its life cycle to the changing parameters of the external environment.

4. Reliability – the existence of an information system in the event of failure of its individual components (information resources), the persistence of system parameter values for a certain period.

5. Interactivity – interaction with the external environment and the "response" variability of information systems.

There are a number of system properties of information systems, such as:

1. Integrativity – the presence of system — forming, system — preserving factors.

2. Equifinality – the ability of information systems to achieve states that do not depend on the initial conditions and are determined only by the parameters of the system.

3. Heredity.

4. Opportunity for development.

5. Self — organization, etc.

For full — fledged work or maintaining a minimum set of critical information functions, an information system must have a well — defined margin of resistance to external destabilizing influences from the external environment (information space), due, in turn, to influences from society, the state, commercial structures, etc. . Both the entire information system and its individual elements can be subject to various destabilizing information impacts, attacks, for example, the removal of individual materials from Internet websites, the destruction or shutdown of information servers, the publication of documents that distort the original information system in a certain direction , or the creation of a new information system that can reduce the relevance or simply destroy the original system.

It is clear that for full — fledged work and preservation of a minimum set of critically important informing functions, an information system must have a well — defined margin of resistance to external destabilizing influences. At the same time, the violation of the integrity of the information system against the background of a decrease in the relevance of its components entails disorganization, a simultaneous loss of flexibility – a decrease in survivability and a violation of integrity, that is, the loss of the most important functions of information systems (Fig. 1) [7].

Figure 1. Model of information system homeostasis

The problems of survivability of information systems require not only expert, high — quality solutions, but also the use of approaches based on mathematical and computer modeling. For example, one can single out the tasks of modeling information operations, which are inextricably linked with the so — called social modeling.

Despite the fact that social modeling is an interdisciplinary direction, it is it that is currently widely used in solving problems in the field of information security [3].

1.2. Information space modeling

An information space is a set of elements (documents) related in meaning that form information systems – clusters of related documents. At the same time, it retains its stable patterns throughout its existence. Numerous studies have shown that the parameters of the frequency and rank distributions of documents in many information systems remain the same and are determined by parameters that depend on the content and subject matter of information. In this regard, S.A. Ivanov [8] noted that "information space is a documentary environment in which cluster structures of scientific publications in periodicals are formed, which are fractals." Information systems reflect communication processes in the information space in their subject area, the emergence of new topics is accompanied by the emergence of new fractal arrays in the information space.

Like many other complex systems, the information space can be represented as a communication environment – in the form of a system with a complex of links between information sources and converters that influence each other depending on the level of perception of the individual information messages generated and transformed by them.

At the same time, on the one hand, classical information theory is quite suitable for modeling information sources and converters, i.e. mathematical theory of communication, developed by Shannon in the 40s of the twentieth century and significantly supplemented and expanded in subsequent years by the works of N. Viner, V.A. Kotelnikov and A.N. Kolmogorov. However, the classical information theory does not take into account the interaction between information sources and converters, which, on the other hand, fits into the ideology of the modern theory of complex systems and is implemented, for example, using multi — agent modeling.

The idea of multi — agent (*Agent — based)*modeling arose in the middle of the twentieth century. In accordance with it, an agent is some abstract entity that has activity, autonomous behavior, can make decisions in accordance with a certain set of rules, can interact with the environment and other agents, and can also evolve itself.

Agent — based modeling (or multi — agent modeling) is a computer — based methodology that allows researchers to design, analyze, and explore formal models represented as some “artificial worlds” inhabited by agents that interact with each other based on predetermined relative simple rules. Multi — agent models vary considerably in their rules, but they all consist of sets of autonominal, myopic (perceiving information only from some local area) units. Although the individual actions of agents are completely determined at the micro level, when a large number of such agents are functioning, their interactions form a macro level that is not always predictable. At the same time, today multi — agent models are widely used to analyze decentralized systems, the regularities of the functioning dynamics of which have not been sufficiently studied. Multi — agent models are used to get an idea of the general behavior of such systems, to identify the rules for the functioning of systems based on assumptions about the individual behavior of their individual components (agents).

The goal of multi — agent modeling can be formulated as the creation of computer microworlds in which agents interact, reacting to conditions from their environment and making changes.

Within the framework of the simplest model of the information space, an information block (document) can be generated by an agent A_1, transmitted to a certain set of agents $A_2, ..., A_n$, which transform the information received, distorting it (accidentally or intentionally, depending on the informational influence of other agents), and transmit it further in the form of their own information blocks . At the same time, blocks similar in information content can be grouped into clusters – information systems. Below, one of the simplest multi — agent models that implement the dissemination of news in the information space, the information diffusion model, will be considered in detail.

According to the definition of K. Langton [9], the modeling of complex adaptive multi — agent systems is often based on the following principles:

— the model consists of a population of simple agents;

— there is no single agent (center) that directs the rest of the agents;

— each agent considers in detail the ways in which a simple reaction to local changes in the environment is carried out, including contacts with other agents;

— there is no single rule in the system that would describe global behavior.

In accordance with these principles, any behavior at the level above the individual is emergent, generated by the interactions of local agents, i.e. "simple rules can cause complex behaviors and structures."

1.3. Multi — agent systems

Multi — agent models are, by definition, decentralized. At the same time, the complex global behavior of the system is the result of the activity of a large number of agents, each of which follows simple rules. Multi — agent models make it possible to study a fairly wide range of problems for which rigorous analytical methods are ineffective.

1.3.1. Artificial societies

In many cases, multi — agent models are implemented in the form of so — called "artificial societies" (eng. – *Artificial Societies)*as computer models of a real society. Thus, artificial society is a subclass of multi — agent models. Artificial society agents behave autonomously: they make decisions, act independently and interact with other agents. The literature describes such areas of application of agency models as, for example, choosing a jurisdiction to live in, voting for political parties, economic policy, creating or modifying jurisdictions, making decisions about participation in communities, etc.

J. Epstein [10] identified the following main characteristics of multi — agent models:

— inhomogeneity. Agents are somewhat different from each other, which fundamentally distinguishes these models from most individual — oriented models;

— autonomy – agents act independently of each other;

— the space in which the agents operate is set or described in advance in an explicit form;

— presence of local interactions;

— limited rationality inherent in agents;

— the presence of the dynamics of the system, except for cases of achieving equilibrium.

Thanks to the interaction between agents, social processes and procedures can be modeled, which is why an artificial society can be considered as a certain class of oriented models.

1.3.2. The heat bug model

Modeling agents and multi — agent systems without the use of modern modeling tools (for example, AnyLogic) is a fairly simple procedure. So, in [11, 12], as an example, a simple model of collective behavior, called "heat bugs" (English – Heat Bugs)[13], is considered. In a discrete environment, divided into cells, "beetles" move, releasing heat, which spreads in the environment. Each individual beetle has its own "ideal" temperature of the environment in which it prefers to be, and it has a sensor with which it can determine in which direction the temperature of the environment is closer to its "ideal" temperature. This allows the beetle to find the direction in which it must move in order to reach the cell with a temperature that suits it. The environment has the following characteristics:

— heat spreads evenly in all directions at a speed proportional to the temperature difference in neighboring cells;

— heat "decreases" in each cell in proportion to the amount of heat that the cell possesses.

The "world of bugs" is dynamic and difficult to predict. For example, even being in a cage with an “ideal” temperature for it, the beetle heats it up, releasing heat, so over time it can move to a more comfortable cage for it, as the state of the environment changes.

Animation of the "world of heat bugs" model is carried out by marking the temperature in each cell with a certain color. As a result of the simulation, it can be seen that the beetles gather in groups, warming each other. The model allows you to trace the behavior of each particular beetle, determine its parameters, coordinates, and ideal temperature. Experiments with this model show how changes in local parameters affect the global behavior and survivability of the entire system.

1.3.3. Natural computing

Recently, within the framework of the concept of multi — agent modeling, the so — called *Natural Computing*, which combines mathematical methods that contain decision — making principles similar to mechanisms implemented in nature, have been intensively developed. Imitation of the self — organization of an ant colony (or termite colony) forms the basis of the so — called ant optimization algorithms [14], one of the promising methods of natural computing, in which an ant colony is considered as a multi — agent system, where each agent functions autonomously according to fairly simple rules. In contrast to the primitive behavior of agents, the behavior of the entire system turns out to be very complex, close to reasonable. Ant colony algorithms are based on simulating the self — organization of social insects through the use of dynamic mechanisms that ensure that the system achieves a global goal as a result of low — level interaction of its elements, provided that the elements of the system use only local information, excluding centralized control. At present, good results have been obtained in ant colony optimization of such complex combinatorial problems as the traveling salesman problem, optimization of transport routes, graph coloring, optimization of network graphs, scheduling, optimization of processes in distributed non — stationary systems, for example, traffic in telecommunication networks.

1.3.4. Group Preference Model

As an illustration of the application of multi — agent systems, consider another model – Axelrod and Hammond [15]. In accordance with this model, the preferences of groups of people were studied. At the same time, it was first assumed that the groups differ only in ethnicity. However, the constructed model can also take into account any other types of difference in which individual group membership is visible and stable.

In the Axelrod — Hammond model, an agent is an individual. Each agent is "colored", which can be interpreted as his ethnicity or other sign of group membership e.

Each agent also has a two — part strategy. The first part of the strategy determines whether the agent cooperates (or not) with a neighbor that has the same color. The second part of the agent's strategy determines whether the agent cooperates with a neighbor that has a different color. As with all multi — agent models, the rules for the interaction of the agent are established first, and then computer simulations are used to trace the evolutionary history. The original goal of the project was to understand the conditions under which the populace would end up putting people in power who would only care for members of their own group and refuse to provide assistance to members of other groups.

Since the purpose of the model was heuristic, the main design criterion was simple. The description of the model is very simple. We consider a system of cellular automata – a field of 50 $\times$50 cells, each of which has 4 neighbors – a von Neumann neighborhood. In this model, agents can also be interpreted as small groups of families such as villages with a single ethnicity. 2000 iteration steps are considered, during which the following actions can be performed:

1. Immigration. An immigrant agent with random genes appears on a random empty lot.

The agent has three genes:

— sign: one of four colors ;
— choice when meeting an agent with the same color: help or not ;
— choice when meeting an agent with a different color: help or not.

2. Interaction. Each agent starts the period with a replay potential (PTR) 12%. Each adjacent pair of agents solves the "one — step prisoner's dilemma":

— if the agent provides assistance: PTR decreases by 1%;
— if the agent receives assistance: PTR increases by 3%.

3. Reproduction. Randomly, each agent with probability PTR reproduces in an adjacent empty cell, if available, with mutation/gene = 0.5%

4. Death. Each agent has a 10 % chance to die.

One of the results obtained from the analysis of the work of the model showed that, most often, the ability of agents to distinguish between a circle of people with common interests and groups leads to a higher level of cooperation than if the agents were “color blind”.

When applying the model to solve real problems, Axelrod and Hammond initialized its space, taking into account the geographical and historical data of Central Asia. For this, a map of the distribution of ethnic groups was used, which were "attached" to the model map. The model was used in places where ethnic conflicts were expected. Since the model is very simple, a high degree of accuracy in its "predictions" was not expected, but it is of some interest in identifying some potential conflict areas.

Known adaptation of the above model to study certain aspects of modernity, for example, the dynamics of growth in the number of media. This could be done by adding to the model a mechanism that allows taking into account the influence of media of different range on the agent and its neighbors. With this model, it can be seen whether the provincial media will become strong enough to challenge the dominance of the national media.

The proposed model could also explore the potential effects of multiple and cross — cultural linkages such as rich/poor, urban/rural, Christian/ Muslim differences.

1.4. Simulation models

Simulation models, including multi — agent ones, are convenient for asking “what if?” questions. [16]. In fact, due to the fact that such models usually contain a probabilistic element, by running the simulation many times, it is often possible to obtain a complete distribution of the situations that occur.

Recently, in connection with the rapid development of computer technologies, the class of so — called simulation models is important and promising from the point of view of practical application of mathematical modeling. Such a model is an algorithm by which a computer generates data sets that describe given characteristics of a real system of interest. At the same time, the operations performed by the machine have nothing to do with the nature and properties of the system under study. It should be noted that the very fact of finding out the possibility of simulation modeling is a considerable achievement of modern science. Indeed, it turns out that the structure of a real process to a certain extent does not depend on its nature and, so to speak, on its material basis. The numbers obtained as a result of manipulating other numbers according to certain abstract rules can exactly correspond to numbers that describe specific processes occurring in the real world.

Of course, when developing a simulation model, the properties of the phenomenon under study are taken into account, but not at the level of internal mechanisms that are either unknown or too complicated for explicit use, but at the level of general characteristics of the corresponding processes.

In terms of practical application, simulation models are good because they allow you to conduct so — called machine experiments, the purpose of which is to study the change in the behavior of the object of study depending on changes in internal parameters and (and) external conditions. Such techniques make it possible to determine the course of development of events that, for one reason or another, cannot be realized in real life. For example, in order to analyze the distribution of refugee flows in the event of California flooding, such a catastrophe should not actually occur. Simulation modeling (in the presence of satisfactory models) in such cases makes it possible to obtain data at a quite acceptable level of accuracy.

The construction of simulation models is a rather complex task that requires , in addition to knowledge of the subject area, also high professionalism in the field of programming. However, if successful, the results will justify the costs.

1.5. Individual — centric models

One of the varieties of multi — agent modeling – individual — based modeling (English – *Individual — Based Modeling)*is a relatively young direction. The first works in this area appeared in the 80s of the XX century. Individual — oriented modeling can be considered, on the one hand, as a kind of multi — agent modeling, and on the other hand, as a logical development of the quantitative ecology of populations, which was initiated in the works of J. Forrester (in particular, his model "World Dynamics" [17]), who was the first to build and calculate high — dimensional models using computer — assisted simulations.

Individual — oriented models are multi — agent simulation models in which the integral characteristics of a population are the result of many local interactions of agents (individuals). The model consists of a description of the environment in which interactions take place, and a set of individuals for whom the rules of behavior and characteristic parameters are defined, which can also change in the course of evolution. An individual — oriented approach to modeling involves the creation of simulation models that take into account some properties of individual individuals and their local interaction in order to build integral models of entire populations formed from many individuals. The individual within these models is considered as a unique, discrete unit, which has a certain set of characteristics that change during the life cycle. Models based on this approach are built from the bottom up, starting with the "parts" of the system (individuals), eventually describing the entire population. The goal of research is often to understand how the properties of a system arise from interactions between parts [18].

Building a model at the level of describing an individual provides a number of advantages, such as transparency in relation to objective natural mechanisms, the ability to describe the object under study with a high degree of detail, and extract more useful information from the simulation results.

Individual — centric modeling is one of the branches of complex systems theory, which aims to provide researchers with modeling tools for solving problems that cannot be considered by traditional methods.

The first individual — oriented models were related exclusively to specific tasks from the fields of biology and ecology and were not of a systemic nature. As will be shown below, they are quite logically transferred to the field of study of social processes and procedures.

Of course, the construction of an integral model of the system based on an approximate description of the rules of behavior of an individual may turn out to be very far from reality, but in this case, much depends on the level of description of these rules, properties
individual individuals and expected population dynamics. Individual — based modeling provides a number of advantages such as the simplicity of describing individuals and their local interactions, the ability to refine these descriptions during modeling, as well as the transparency of the “rules — model — reality” feedback.

One of the most popular models in the framework of individual — oriented modeling is the so — called "sugar model" of an artificial society by J. Epstein and R. Axtel [19].

Despite its simplicity, this model is a powerful tool for analyzing information systems, in particular, such a property as survivability. Today, this model has many modifications, even the creators themselves have found a way to describe the same model in the formalism of cellular automata. Here is a description of the original version of the "sugar model".

The space in this model is a two — dimensional grid with equal cells. At every moment of time t there is a constant finite number of agents located in space. At the point in time t each cell (x, y) can contain an agent a ($a_t(x, y) = a$), i.e. agent a is in a cell (x, y), or does not contain an agent ($a_t(x, y) = \varnothing$). The amount of "sugar" in the cell (x, y) at the moment of time t is a certain value depending on x and y.

The agent appears on the field (is born) with two parameters: vision (the number of cells in the grid that he can see) and the metabolic rate (the amount of sugar he eats per unit of time to survive). The agent can tolerate any amount of sugar. If the agent does not have the proper amount of sugar to eat, then he dies. Simultaneously with the "death" of one agent, another agent is "born" with randomly chosen parameters and location in space. Thus, the size of the entire population of agents is a constant value. The rules for the behavior of agents are as follows:

— the agent studies the neighborhood of his vision (four or eight directions of the grid) and determines the free cell that has the largest amount of sugar;

— after that, the agent moves into this cell and collects all the sugar.

On fig. 2. different stages of the evolution of the sugar model are presented.

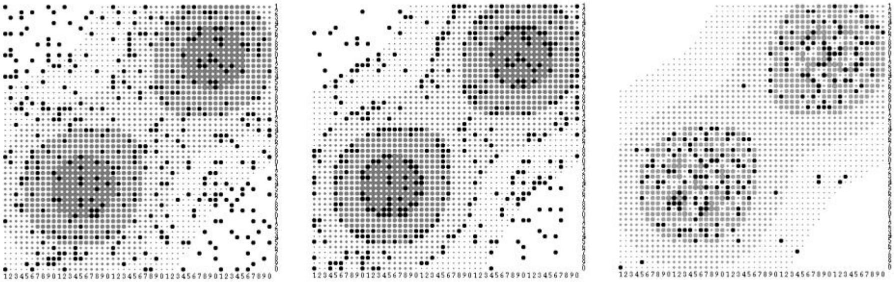


Figure. 2. Evolution of the simplest sugar model

Based on the sugar model, the authors obtained results that are consistent with normal social behavior. Already on the simplest version of the model without explicit interactions between agents, the authors explore the possibility of combating population concentration, in particular, in connection with the problem of environmental pollution. For this purpose, environmental pollution as a result of the collection and consumption of sugar is introduced into the model i.e. In this case, each cell contains sugar and some level of pollution. According to the new rules, the agent moves to a free cell, where the ratio (sugar/level)of contamination is maximum. The introduction of these additional conditions leads to fundamental changes in the evolution of the model: the number of deaths increases, and the distribution of the population becomes more uniform.

In later modifications of the sugar model, different types of interactions between agents are considered, as well as other complications. This makes it possible to analyze a wider range of social processes and procedures.

The Sugar Model explores the following questions:

— distribution of wealth among agents;

— distribution of agents by age;

— migration of agents;

— introduction of new model properties into the model, for example, environmental pollution, and the corresponding modification of the rules;

— introduction of new properties of agents, for example, the division of agents by sex;

— changing the rules for the birth of new agents;

— the introduction of inheritance rules, for example, when the wealth of a deceased agent is evenly distributed among his descendants;

— the introduction of many products, such as sugar and honey;

— introduction of rules for the exchange of products between agents (trade), etc.

Individual — oriented modeling makes it possible to describe the following properties of the object being modeled (especially important for modeling social processes):

— consideration of spatial aspects;

— taking into account the impact of material resources and the influence of the media;

— taking into account social aspects and individual characteristics;

— taking into account the level of detail.

Accounting for spatial aspects. Individual — oriented modeling covers spatially distributed models (English – *Spatially — explicit Models*), in which each individual is associated with a certain position in space.

For a full — fledged modeling of electoral fields, as a rule, spatial distribution is desirable, taking into account movements in space. For example, when modeling possible election fraud by absentee ballots.

Accounting for the impact of material resources and the influence of the media. As it is known, electoral processes essentially depend on the material means and influence of mass media used during election campaigns.

Models that take into account information flows are called informational. In these models, the description of the change in the state of individuals is based on rules based on the analysis of information flows (including, for example, the state of neighboring individuals).

Accounting for social aspects and individual characteristics . For modeling electoral processes, it is especially important to describe the entire system at the level of individual individuals.

For electoral processes, social mechanisms play a very important role, so electoral fields are typical social networks. If the model is aimed at studying social mechanisms, requires taking into account individual differences and training individuals , this subclass of models should be chosen.

As is known, individual variability is the fundamental principle of evolution. At the same time, taking into account this factor often significantly complicates the models; therefore, it is not taken into account, in particular, in the method of cellular automata considered below.

Level of detail. The properties of the model essentially depend on its spatiotemporal scale. The models also differ in the number of individuals considered. The amount of computation directly depends on the scale of the task. This fact must be taken into account when choosing the scale of the model and its implementation.

It should be noted that individual — based models require more computation than analytical models. At the same time, for many areas, including the study of information processes and systems, the development of individual — oriented models is justified due to the fact that:

— data of real observations of the studied parameters are often not enough to identify the analytical model;

— it is necessary to take into account spatial aspects;

— it is necessary to take into account the social mechanisms of the population, individual differences of the individual, learning.

1.6. Models of cellular automata

Multi — agent modeling also covers the concept of cellular automata – mathematical objects that represent a discrete dynamic system. Cellular automata were first introduced by von Neumann in the 1940s as a formal apparatus for studying complex distributed systems (see [20]).

Easy — to — implement models of cellular automaton systems often shed light on problems that are difficult to analyze by other methods.

Cellular automata are useful discrete models for branches of dynamical systems theory that study characteristic collective phenomena.

Models of cellular automata systems are most often implemented using a planar grid of square cells. Different cell states are represented on the computer monitor by different colors . The evolution of the system of cellular automata is demonstrated as a sequence of changes in the states of individual cells, which, in accordance with certain rules, are updated at each time step of the system of cellular automata. The corresponding rules take into account the state of each cell, as well as its surroundings. There are different types of neighborhoods in cellular automaton models.

The discreteness of the model, more precisely, the ability to represent the model in a discrete form, is currently one of the significant advantages, since it opens up wide possibilities for using computer technologies. The evolution of a system of cellular automata can be represented as a discrete sequence of steps, with changes at each step determined by changes at the previous one. However, processes that are often continuous in nature are quite acceptable approximated by suitable discrete constructions, but for this, the model, as a rule, must contain a large number of discrete elements, and the evolution must take place over a large number of cycles.

The main advantage of cellular automata is their absolute compatibility with algorithmic methods for solving problems. Therefore, in recent years, due to the rapid development of computer technology, they are widely used in various fields of science and technology.

Cellular automata are essentially spatially non — mobile discrete individually oriented models. In the traditional system of cellular automata, all cells are equal (the space is homogeneous), while in the individual — oriented system, in addition to describing the cells, there is the concept of an individual who can occupy different cells (or several different individuals can occupy one cell). Thus, in a cellular automaton, the cells change their state synchronously, and the simulation cycle is a cell enumeration. In individual models, the cycle may consist of iterating over individuals , i.e. _ in a cellular automaton, modeling is based on dividing space into homogeneous sections; in individual — oriented models, entities are described that change their position in space. Of course, cells in a cellular automaton can be in different states, and by defining complex states, it is possible to model the presence of individuals in cells and their movement between cells. But this is possible only under significant restrictions .

A cellular automaton is a discrete dynamic system, a collection of identical cells interconnected in a certain way. All cells form a network (lattice) of cellular automata. The state of each cell is determined by the state of the cells included in its local neighborhood and called nearest neighbors [21]. The neighborhood of a finite automaton with a number j is the set of its nearest neighbors.

The state of the j-th cellular automaton at the moment of time $t+1$ is defined as follows:

$$y_j(t+1) = F\left(y_j(t), O(j), t\right),$$

where F *is* some rule that can be expressed, for example, in the language of Boolean algebra. In many problems, it is considered that the element itself belongs to its nearest neighbors, i.e. $y_j \in O(j)$. In this case, the formula is simplified: $y_j(t+1) = F\left(O(j), t\right)$. Cellular automata in the traditional sense satisfy the following rules:

— the change in the values of all cells occurs simultaneously (the unit of measurement is a cycle);

— the network of cellular automata is homogeneous, i.e. the rules for changing states are the same for all cells;

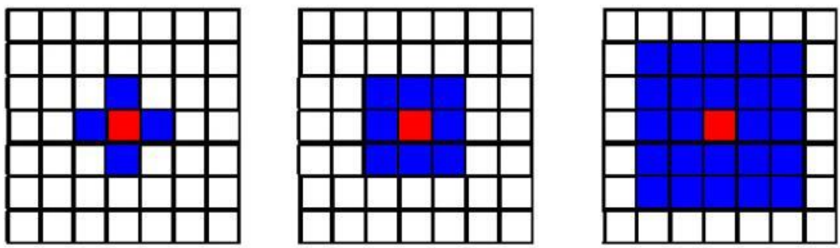
— a cell can be affected only by cells from its local neighborhood;

— the set of cell states is finite.

Theoretically, cellular automata can have any dimension, but one — dimensional and two — dimensional systems of cellular automata are most often considered.

The model considered here is two — dimensional, so further formalism will refer to this case. In a two — dimensional cellular automaton, the lattice is realized by a two — dimensional array. Therefore, in this case it is convenient to pass to two indices, which is quite correct for finite lattices.

In the case of a two — dimensional lattice whose elements are squares, the nearest neighbors included in the neighborhood of the element $y_{i,j}$ can be considered either only the elements located up and down and to the left and right of it (the so — called von Neumann neighborhood: $y_{i-1,j}, y_{i,j-1}, y_{i,j}, y_{i,j+1}, y_{i+1,j})$, or the diagonal elements added to them (G. Moore neighborhood): $y_{i-1,j-1}, y_{i-1,j}, y_{i-1,j+1}, y_{i,j-1}, y_{i,j}, y_{i,j+1}, y_{i+1,j-1}, y_{i+1,j-1}, y_{i+1,j}, y_{i+1,j+1})$. In the Moore model, each cell has eight neighbors, in the extended model, twenty — four (Fig. 3).



Neighborhood: von Neumann *mura* *Extended Moore neighborhood*

Figure. 3. Neighborhoods of two — dimensional cellular automata

To eliminate edge effects, the lattice topologically "folds into a torus" (Fig. 4), i.e. The first line is considered to be a continuation of the last line, and the last line is considered to precede the first. The same applies to columns.

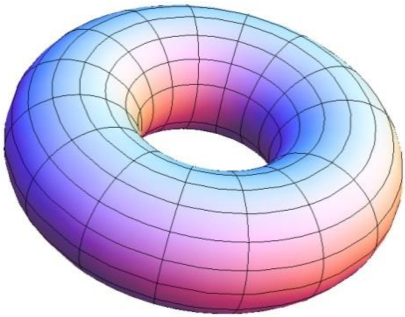


Figure. 4. Rectangular area folded into a torus

This allows you to determine the overall ratio of the value of the cell in the step $t+1$ compared to the step t [22–24]:

$$y_{i,j}(t+1) = F(y_{i-1,j-1}(t), y_{i-1,j}(t), y_{i-1,j+1}(t),$$
$$y_{i,j-1}(t), y_{i,j}(t), y_{i,j+1}(t), y_{i+1,j-1}(t), y_{i+1,j-1}(t), y_{i+1,j}(t), y_{i+1,j+1}(t)) \,.$$

S. Wolfram [21] , classifying various cellular automata, singled out those whose dynamics essentially depend on the initial state. By selecting various initial states, one can obtain a wide variety of configurations and types of behavior. It is precisely such systems that include a classic example – the game "Life", invented by J. Conway and known to a wide circle of readers thanks to the publication in the book by M. Gardner [25] .

Among cellular automata, there are known ones whose dynamics essentially depends on the initial state. By selecting various initial states, one can obtain various configurations and types of circulation. It is to such systems that a classic example belongs – the game "Life", invented by J. Conway.

The rules of the variant of the game "Life" are as follows. The cell is in one of two states – living and non — living (black and white). If there are less than two or more than three black cells in the neighborhood of a cell, then at the next step it is painted white (it dies). If the neighbors of a cell are exactly three black cells, then at the next step it is painted black (is born).

The behavior of systems of cellular automata in the general case can be reduced to the dynamics of complex nonlinear systems [26], however, this is a very difficult scientific problem for each specific system of cellular automata.

Recently, numerous attempts have been made to typify systems of cellular automata. Here is one of the most successful ones.

The systems of cellular automata of the first type include such systems that, regardless of the initial state, perform a transition to a homogeneous state in a finite number of steps – all automata are in a state of rest.

In the process of evolution of cellular automata of the second type, the system comes to localized stationary or periodic solutions.

The results of the activity of the system of cellular automata of the third type are aperiodic, i.e. these systems exhibit chaotic behavior.

The dynamics of systems of cellular automata of the fourth type essentially depends on the initial state. By selecting various initial states, one can obtain a wide variety of configurations.

The literature contains numerous examples of the use of cellular automata models for solving applied problems in the analysis and modeling of social processes and procedures. The works of T. Brown [27] consider a number of contextual models of the electoral process, in which it is assumed that the electoral preferences of an individual are determined by the attitudes of his immediate environment.

In one of the "binary" models, it is assumed that an individual decides to vote at the moment $t+1$ for one of two alternatives: for Republicans or Democrats in accordance with the rules of a simple majority of his environment – the nearest neighbors. If out of nine people five or more support the Democrats, then the individual also votes for the Democrats. If the majority is Republican, then the individual also shares the view of the majority. In this case, the cellular automaton has two states: 1 – voting for the Republicans; 0 – vote for the Democrats.

Some examples of cellular automata used in sociological problems are given in [23, 24]. In particular, a model of the process of racial segregation in the choice of place of residence is described [28].

In this example, it is assumed that each racial group prefers to have a certain percentage of neighbors with the same skin color. If this condition is not met, then the family moves to the nearest house where the percentage of neighbors is acceptable. In [28], a finite automaton model with simple rules and a Moore neighborhood was used. The constructed model quite realistically described the process of dividing the region into several racially homogeneous regions.

A similar model was proposed in [29], within which such rules for the interaction of the existing two types of agents are implemented. The agent prefers to live surrounded by his own kind, i.e. moves to an area where there are more agents like him. As a result, a structure similar to Indian castes is formed.

Cellular automata are also successfully used in modeling the processes of news and innovation dissemination [30].

The article by T. Brown [27] considers the model of electoral process. He believes (with which the authors are in full agreement) that the electoral preferences of an individual are determined by the attitudes of his immediate environment. One model assumes that an individual decides to vote $t+1$ Republican or Democrat at the moment in accordance with the simple majority rule. The views of the individual and his four nearest neighbors at the moment t (the von Neumann neighborhood) are taken into account. The model was studied over a large time horizon, up to 20,000 cycles. It turned out that the party struggle leads to very complex configurations, which essentially depend on the initial distribution.

Another, created back in 1956 the model, known as the Tiebout model [31], considers people located in a finite number of city districts. Within the framework of this model, each resident decides to remain in the jurisdiction of his residence or move to another area in order to maximize the value of his utility function.

The agent makes a decision based on information about jurisdictions: a set of local public goods and the level of taxation. There are many numerical results related to the existence of equilibria of the Tiebout model and its optimal properties.

To study the problems of survivability of information systems as complex multi — parameter systems, the parameters of which are still poorly understood, the most suitable technique is mathematical modeling. The life cycle of information systems can be described, for example, by the information diffusion model [32]. Recall that in the natural sciences, diffusion is understood as the mutual penetration of contacting substances into each other, caused, for example, by the thermal motion of their particles. Information diffusion processes, like diffusion processes in physics, are quite accurately modeled using cellular automata methods.

Cellular automata are useful discrete models for the study of dynamical systems. The discreteness of the model, or rather, the ability to represent the model in a discrete form, can be considered an important advantage, since it opens up wide possibilities for using computer technologies.

The model of information diffusion, which we will consider further, is two — dimensional; therefore, the entire system of cellular automata is described by a two — dimensional array. Within the framework of this model , which refers to the dissemination of news in the information space, the Moore neighborhood [20] and probabilistic rules for the distribution of news on a given topic are applied.

It is assumed that the cell can be in one of three states: 1 – "breaking news" (the cell is painted black); 2 – news, outdated, but saved as information (gray cell); 3 – the cell does not have information transmitted by the news message (the cell is white, the information has not reached or has already been forgotten). The rules for the development of the infoplot are as follows:

— initially, the entire field consists of white cells, with the exception of one – black, which was the first to "accept" the news;

a white cell can only be repainted black or remain white (it can receive news or remain "in the dark");

— the white cell is repainted if the following condition is met: $Cpm > 1$, where p is a pseudo — random value ($0 < p < 1$), m is the number of black cells in the neighborhood, C is a constant ($C = 1.5$ for $m = 1$; $C = 1$ at $m \neq 1$);

— if the cell is black, and around it are exclusively black and gray, then it is repainted in gray (the news becomes obsolete, but is saved as information);

— if the cell is gray, and around it are exclusively gray and black, then it is repainted in white (the information is forgotten when they are well known).

The described system of cellular automata quite realistically reflects the process of infoplot development (Fig. 5). On a field of size 40 ×40 (dimensions were chosen solely for the purpose of clarity), the state of the system of cellular automata is completely stabilized in a limited number of cycles, i.e. in practice, the process is convergent.

Typical dependences of the number of cells (a sequence of the number of cells of the same type) in different states, depending on the iteration steps, are shown in Fig. 6.

When analyzing the above graphs, one should pay attention to the following features: 1 – the total number of cells in all three states at each iteration step is constant and equal to the field size ; 2 – upon stabilization of cellular automata, the ratio of the number of gray, white and black cells is approximately 0.75 : 0.25 : 0; there is a point of intersection of the curves defined by all three sequences at the level of 33% each.

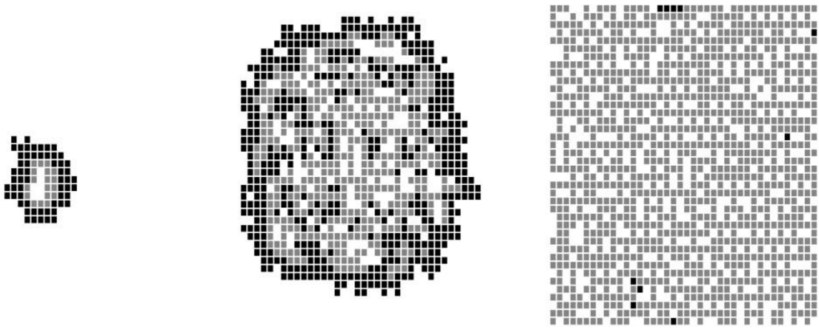


Figure 5. States of evolution of a system of cellular automata

It is black cells that form the actual infoplot, the dynamics of which is shown in Fig. 6.

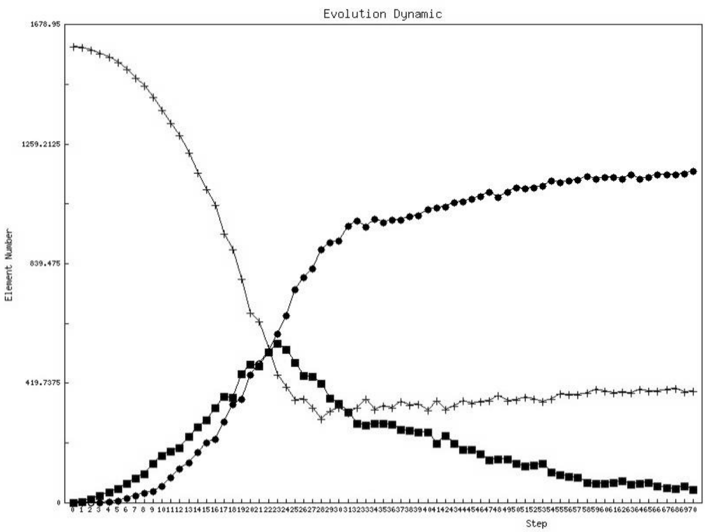


Figure 6. Distribution of cells depending on the cycle systems of cellular automata: white cells – (+) ; *gray cells –* (•) ; *black cells* — (■)

, white x_w and black x_b cells on the step of evolution of the system of cellular automata obtained as a result of analytical modeling are expressed by the formulas: x_g

$$x_g = \frac{0,75}{1+e^{-0,15(t-30)}},$$

$$x_w = 1 - \frac{0,75}{1+e^{-0,25(t-20)}},$$

$$x_b = 0,75\left(\frac{1}{1+e^{-0,25(t-20)}} - \frac{1}{1+e^{-0,15(t-30)}}\right).$$

Life cycle of infoplots The same, like many other systems, can be described using two more large classes of models: Boolean and Markov.

It should be noted that the dependence of news diffusion obtained as a result of modeling is in good agreement with the actual behavior of thematic information flows on Internet sources, and on local time intervals with traditional models.

2. ANALYSIS AND EVALUATION OF SURVIVABILITY

It is necessary to assess the system's ability to continue normal functioning under conditions of permanent destructive influences and resist them, adapt the functioning algorithms to new conditions and organize functional restoration or ensure functioning during a gradual degradation process, possibly without losing the most significant "critical" information functions ; it is necessary to move from the analysis and assessment of reliability to the analysis and assessment of survivability.

Under survivability, we mean the ability of an information system to maintain and restore the performance of basic functions in a given volume and for a given time in case of a change in the structure of the system and / or algorithms and conditions for its functioning due to adverse effects (I).

In addition to the possibility of "internal" recovery of the system after NV, the survivability of the system is also characterized by the possibility of influencing the external environment in which the system itself operates. This possibility is particularly clearly visible just in the case of information systems.

One of the indicators of the survivability of the system is the margin of survivability (*d* – survivability) – the critical number of defects, reduced by one. By a defect we mean a unit of measurement of the damage caused to the information system of the NV. If we denote by *C* the critical number of defects, then the indicator d — live — honor will be $d = C - 1$.

Critical is the minimum number of defects, the occurrence of which leads to the loss of the information system of its performance (possibility of information impact).

On the other hand, the margin of survivability can be defined as the maximum number of defects that the system can still withstand without loss of performance.

Let $m_i = i$ the i — th combination of defects, in which the system does not lose its performance, then the survivability margin is defined as $m = \max\limits_{(i)} m_i$.

Models of analysis and assessment of survivability can be static and dynamic. In static models, the site of damage to the information system and the intensity of the influence of specific types of NI are specified, a list of elements that can be affected by NI (for example, website pages) is determined, and with the help of the logical health function, an indicator of the quality of the system functioning is found. Dynamic analysis models are simulation models that include: a model for the emergence and development of NI, a model for changing the states of elements of an information system under the influence of NI, and a model for functioning under conditions of changes in the structure and values of system parameters associated with NI.

2.1. Functional survivability

The survivability of systems is analyzed and evaluated at various levels of design, modeling and operation of information systems. During the study of functional survivability, game — theoretic, probabilistic, graph, matrix models can be used .

When studying the functional survivability of information systems, the features of the topology of the network of intercomponent links are taken into account indirectly. It is assumed that information systems provide the necessary connectivity of operable components.

When analyzing the survivability of functioning, the information system is characterized by:

— the purpose of functioning (informing, misinforming, information impact, etc.);

— a set of tasks $Q = \{q_1, ..., q_m\}$, the solution of which is provided with its help;

— a set of components (information resources) $\{S_1, S_2, ..., S_p\}$, which are the components of the system.

During the functioning of an information system, its components can be in one of the states: operable, not operable, partially operable, i.e. efficient, but with a partial decrease (within acceptable limits) of the value of any indicators of the quality of functioning.

On the basis of game — theoretic models, the survivability of systems that operate under conditions of purposeful influence is investigated —

of the enemy, external and internal destructive influences, when it is possible to compensate for emergency situations, flows of failures and failures only at the expense of the internal reserves of the system and the impact on the source of destructive influences.

Probabilistic, graph, matrix models of analysis and assessment of survivability are quite diverse. In each specific case, for different models, taking into account the various goals of functioning, as well as the conditions for the system's performance, it is possible to find quantitative estimates of survivability. The survivability indicators of different systems can be compared if the goals of their functioning coincide.

Quantitative indicators of survivability significantly depend on the parameters that determine the conditions for the operability of the information system. The current level of performance determines the quantity, quality and content of functions, which are summarized by the concept of "the purpose of the system functioning". To ensure the goal of the system functioning, one of the strategies can be applied:

— f — strategy — fault — tolerance strategy (*fault — tolerance*);

— s — strategy – strategy to ensure survivability (*survi — vability*).

In the process of forming f a strategy, it is necessary to determine the set of system states $S^{(f)} = \left\{ s_v^{(f)} \right\}$ in which it is necessary to counteract the threats of disruption of performance, to set options for the distribution of functions between the operable components of information systems in the states of the set $S^{(f)}$.

The fault tolerance strategy is focused on fully compensating foreseeable functional failures and providing performance indicators for systems in these cases.

In the process of forming s strategies for each state of the set, $S^{(f)}$ it is necessary to additionally develop decisions related to the functions of the system: whether or not to narrow down the set of functions that together constitute the goal of functioning; how to do it; whether or not to simplify the algorithm for implementing functions, etc.

The solution option regarding the purpose of the system functioning in the presence of undesirable influences can be one of the following:

1. A set of system functions cannot be changed, all functions must be used, possibly with less efficiency or with deterioration in quality, i.e. in any state of the following $S^{(f)}$ condition must be met:

$$\prod_{i\in I} x(f_i)=1,\ \ x(f_i)=\begin{cases}1, & \text{если}\ \ f_i\ \ \text{выполняется};\\ 0, & \text{если}\ \ f_i\ \ \text{не}\ \ \text{выполняется}.\end{cases}$$

2. In any state from, $S^{(f)}$ some subset of functions must be performed F^* that realize the purpose of the information system functioning, i.e.

$$\prod_{f_i\in F^*} x(f_i)=1.$$

The set of functions F^* depends on the state of the system and the given conditions for functional survivability.

3. In an arbitrary state from $S^{(f)}$ the system must ensure the execution of at least one function from the set F^*, i.e.

$$\sum_{i\in I} x(f_i)\geq 1.$$

The functional survivability of an information system depends on the predetermined purpose of its functioning. The functional survivability of different information systems can be compared if they have the same functioning goals. The assessment of the survivability of the same information system may change in case of a change in the purpose of functioning. At the same time, the parameters that determine the conditions for their performance have an equally significant impact on the quantitative indicators of the survivability of information systems, as the goal of functioning.

When choosing mechanisms for increasing the functional survivability of a particular information system, it is necessary to take into account the purpose of functioning (the set of functions that the system implements), the structure of relationships, and the features of functional components.

Under the purpose of functioning, they have in kind of a concept that is introduced for tenacious and fault — tolerant information systems, but its change is assumed to be possible only for systems with the property of survivability.

The qualitative dependence of the purpose of functioning (the number of information functions performed) on the number of component failures for fault — tolerant and survivable systems is shown in Fig. 1, respectively. 7a and *7b* . _

At the same time, there are tenacious systems for which the dependence under consideration is expressed by a dependence that is something intermediate between the dependences shown in the s. The behavior of such systems (systems with phase transitions) will be considered in detail in the chapter devoted to the critical level of survivability of systems.

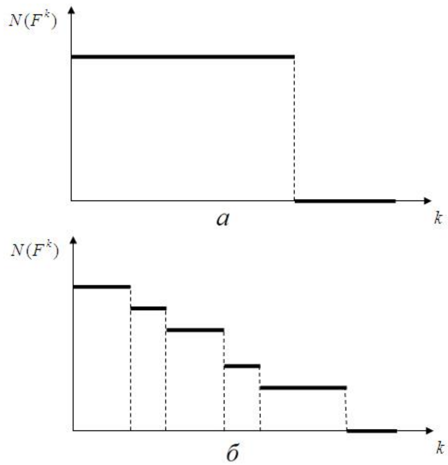


Figure 7. Dependence of the number of functions performed (vertical axis) on the number of component failures (horizontal axis, k)

In the process of analyzing and evaluating the functional survivability of an information system, it is assumed that it is possible to provide the necessary links between individual functional components.

Let us denote the set of information functions implemented by the system as $F = \bigcup_{i \in I} F_i = \{ f_1, f_2, ..., f_n \}$, and the functional component Φ_k can potentially perform a set of functions $\varphi_n : \{1, 2, ..., p\} \to P(F)$, where $P(F)$ is the set of all subsets of F . If $\varphi_n(k) = \{ f_{k_1}, f_{k_2}, ..., f_{k_j} \}, 1 \le k_r \le n,$ then the functional component Φ_k can perform the functions $f_{k_1}, f_{k_2}, ..., f_{k_j}$.

At each specific point in time, the functional component Φ_k is intended to perform a certain set of functions defined in this way: $\varphi_t : \{1, 2, ..., p\} \to P(F)$. If $\varphi_n(k) = \left\{ f_{k_1}, f_{k_2}, ..., f_{k_j} \right\}, 1 \le k_r \le n,$ then the functional component Φ_k can perform the functions $f_{i_1}, f_{i_2}, ..., f_{i_j}$. If $\varphi_t(k) = \varnothing,$ it's Φ_k not working.

Each function $f_i \in F$ is characterized by some performance efficiency c_i (for example, execution time, audience size when submitting information, etc.). The efficiency function for the system can be defined as $\varphi_{ef} : F \times \{1, 2, ..., p\} \times \quad \times P(F) \to C,$ where C is some numerical set; $\varphi_{ef}(f_i, k, \varphi_t(k)) = c_{i_k}$ means that the functional component Φ_k is designed to perform functions $\varphi_t(k) = \left\{ f_{i_1}, f_{i_2}, ..., f_{i_j} \right\}$, then the performance efficiency $f_i \in \left\{ f_{i_1}, f_{i_2}, ..., f_{i_j} \right\}$ is equal to c_{i_k} .

The conditions for achieving the goal of functioning (performing a certain set of functions with a given efficiency) can be defined as follows:

$$\bigcup_{k=1}^{P} \varphi_n(k) \supseteq F;$$

$$\varphi_t(k) \subseteq \varphi_n(k) \quad \forall k = \overline{1, p};$$

$$\sum_{k=1}^{p} \varphi_{ef}(f_i, k, \varphi_t(k)) \ge c_i, \quad \forall i = \overline{1, n}.$$

By functional failure we will understand the impossibility of performing a certain function by a functional component. In this case, the state of the functional component changes and, accordingly, the function changes φ_t. If this leads to a violation of the conditions for achieving the goal of functioning, then the means of ensuring survivability must correct the functioning of the system in such a way that the above conditions are met, possibly by changing the function φ_t. The optimality of the system behavior can be determined by fulfilling the necessary conditions by changing the minimum quantity φ_t, i.e. changes in the number of functional components involved in failure compensation procedures. In this case, the number of compensated functional failures can be the criterion of system survivability (quantitative assessment).

If the components Φ_k are functionally homogeneous, and the purpose of functioning is determined through a given level of system performance, which is ensured due to the presence of an appropriate number of functional functional components:

$$R(\Phi_k, t) \geq R^* = \text{const},$$

where $R(\Phi_k, t)$ is the average number of operable functional components in the system at a point in time $t \geq 0$, R^* is the minimum allowable number of operable functional components, at which the system performance is not less than the required one, then the functional survivability assessment can be the function

$$N(\Phi_k, t) = \overline{\Omega}(\Phi_k, t)/(N\omega),$$

where $\overline{\Omega}(\Phi_k, t)$ is the mathematical expectation of system performance at a point in time $t \geq 0$; $N\omega$, which is the total performance of all functional components.

2.2. Structural survivability

Structural survivability is considered as the possibility of reconstruction, reorganization, reconfiguration during NV, which will allow creating a structure that ensures the performance of a critical subset of functions to achieve the goal of the system functioning.

When considering structural survivability, the topology of the intercomponent communication network and the reliability characteristics of the components are taken into account. The problems associated with the analysis of structural survivability can be reduced to problems of reliability, connectivity of topological structures, depending on the introduction of the concept of "destruction".

Analysis of structural survivability requires the determination of:

— structures to fulfill the purpose of the system functioning at some point in time when undesirable influences on the system occur;

— requirements for certain types of system resources and their relationship;

— requirements for the functionality of the system components;

— features of the nature of undesirable influences or their consequences.

The structural survivability of a system can be assessed under certain assumptions, which make it possible to simplify the assessment task and reduce it to the task of analyzing the connectivity of graphs, estimating the probability of forming a workable structure in the event of undesirable influences, etc.

When considering an information system, it is also necessary to take into account the number of structures in this system that can perform a critical subset of information functions. To calculate this number, it is necessary to isolate such structures, i.e. switching to the language of graph theories and complex networks, to find in network structures a set of components connected in a certain way (not necessarily "strongly connected"), for example, cliques, taking into account that connections are formed due to semantic dependencies. The transition to the consideration of network models should also be based on an assessment of the characteristics of networks, their maximum and minimum sections, flows in these networks, costs, etc.

In the study of structural survivability using graph models, the set of system components $G(V,R)$ is depicted as graph vertices $v \in V$, and the graph edges $r \in R$ correspond to the connections between them. A system that is modeled using a graph is considered to be destroyed if, in the event of removal of vertices or an edge, the graph satisfies one or more of the following conditions:

— the graph consists of at least two components;

— there are no directed paths for certain sets of vertices;

— the number of vertices in the largest component of the graph $G(V,R)$ is less than some predetermined number;

— the shortest path exceeds some given value.

Accordingly, the system is considered viable if these conditions are not met. Structural survivability of systems is usually characterized by various indicators of connectivity. The calculation of indicators such as, for example, the probability of connectivity under conditions of random existence of graph edges, in practice is limited by the computational complexity of such problems. At the same time, using the paths and discontinuities of the graph modeling the system, one can obtain fairly simple limit estimates of the required indicators.

Let's describe the mechanism of reconfiguration of an information system (information cluster) under a destructive effect – the removal of an element (document).

Let $C_1, C_2, ..., C_M$ — various information systems that affect the subjects (objects of influence) $U_1, U_2, ..., U_L$. The information system C_j consists of elements $D_{11}, D_{12}, ..., D_{1J}$.

At the same time, the strength of the influence of each element D_{ij} of the information system C_j on the subject U_k is estimated by the value $f_v(D_{ij}, U_k) =$
$= V(i, j, k) \geq 0.$

On fig. 8, *a* shows a diagram of the influence on the subjects U_1, U_2, U_3 of the information system C_1. At the same time, the subject U_2 also has a non — zero impact from the information system C_2 through the element D_{21}. It should be noted that both between elements from one information system and between separate information systems and their elements there are information links (not shown in Fig. 8).

As a result of a destructive impact by removing an element, D_{13} the information system C_1 loses its influence on the subject U_2 (Fig. 8, *b*).

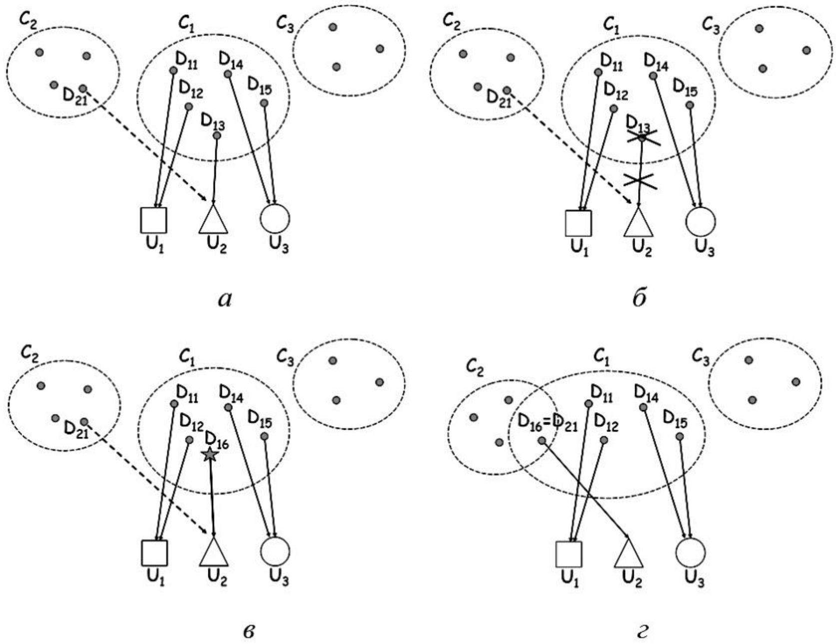


Figure 8. Recovery after a destructive impact to information system C_1

Restoring influence on the subject U_2 is possible by reconfiguring the information system C_1 in three directions:

1) generation of new documents affecting U_2, for example, D_{16} (Fig. 8, *c*);

2) modifying existing messages from C_1, adding content to them that affects U_2;

3) modifying the message D_{21} in the direction of the subject of the information system s C_1, and, thus, combining a part of the information system C_2 with C_1 (Fig. 8, *г*). In real life, this path can be difficult, for example, when the information system C_2 belongs to forces that compete with (or are in antagonistic relations with) the owners C_1.

In any case, it is possible to estimate the cost of creating/reconfiguring an information system.

Based on the fact that the overall assessment of the impact of the information system C_i is determined by the formula:

$$W(i) = \sum_{k=1}^{L} \sum_{j=1}^{|C_i|} f_v(D_{ij}, k),$$

where $|C_i|$ is the number of elements in the information system C_i, with the total cost of creating this information system is

$$S(i) = \sum_{j=1}^{|C_i|} E(D_{ij}),$$

where $E(D_{ij})$ is the cost of creating an element D_{ij} of the information system C_i, then one of the ways to determine the effectiveness of the information system is to evaluate the expression

$$Q = \frac{\sum_{k=1}^{L} \sum_{j=1}^{|C_i|} f_v(D_{ij}, k)}{\sum_{j=1}^{|C_i|} E(D_{ij})}.$$

An important group of characteristics of structural survivability are the so — called "survivability measures". Determining them, they come out of
the assumption that the adversary, knowing the structure of the network, is trying to disrupt its functioning. The survivability of a system is considered high if it is necessary to destroy a significant number of nodes and/or ribs in order to significantly impair or interrupt its operation.

In the mathematical theory of graphs, vitality indicators are interpreted as quantitative measures of connectivity for the graph structure: the minimum section, nodal connectivity, generalized connectivity, path length, etc.

For the design of information systems, the problem of analyzing structural survivability can be formulated as the problem of estimating the value of the maximum flow that can be transmitted in the network in the event of failure of its elements and lowering to an acceptable level of quality of operation.

For example, the structural survivability of an MPLS (*Multiprotocol Label Switching)*network [2] is evaluated using the following algorithm:

1. Calculate the total value of the flow in the fail — safe state for all classes of service;

2. Simulate various failure states and calculate the probabilities of these states;

3. For each failure condition, calculate: the relative reduction in overall throughput; optimal flow; maximum flow; bandwidth reserve, etc.

Structural survivability — related automatic reconfiguration and reloading, for example, in the event of DoS attacks, are today de facto mandatory functions when organizing web and proxy servers, in particular under the control of the Apache HTTP server (http://apache.org/). In particular, there are plug — ins for this web server that make changes to its configuration and ensure that it restarts when DoS attacks are detected. Currently, for another HTTP server – nginx (http://sysoev.ru/nginx/) – such tools are mandatory components. The ngx_http_limit_req_module of this system allows you to limit the number of requests for a given session or, as a special case, from one address. The ngx_http_limit_zone_module module allows you to limit the number of simultaneous connections for a given session or, as a special case, from one address. In December 2009, nginx was used by 4% of the most visited sites in the world.

The restart of the Squid proxy server (http://www.squid.org/) is handled by a special memory — resident program (the so — called daemon) samsdaemon. This program checks at specified intervals to see if a reconfiguration is necessary. For example, if the user has exceeded the limit of his authority, it is possible to automatically enter his data into the corresponding configuration file of the proxy server. Thus, the reconfiguration of the Squid proxy is done by modifying the squid.conf file and then signaling the reconfiguration.

2.3. Game — theoretic approach

In game — theoretic models, the survivability of information systems can be viewed from the point of view of the exchange of a certain amount of spent V resources and a certain amount of consumed W resources. In the process of quantifying the survivability of systems that are described by models of this type, information systems correspond to the promotion of goods on the market in the economic environment. At the same time, the content of — exchanges is introduced and specified (V,W), and these exchanges are optimized. The survivability of information systems can be assessed by the presence of vital elements in them, which are determined by expert assessments.

Consider a system A in interaction with a system B.

Each system in time can translate itself into different states, changing its structure and behavior. Let us assume that the goal $\underline{A}$ of the system A is its desire to achieve certain states that are preferable for it, the goal $\underline{B}$ systems B is the same striving to achieve states that are preferable to her. The expediency of the structure $|A|(|B|)$ and purposefulness of the behavior $\bar{A}(\bar{B})$ of the system $A(B)$ is assessed by the efficiency with which the system achieves the goal $\underline{A}(\underline{B})$. So the system $A(B)$ when fixing its goal, it can be characterized by two factors:

$$A=\left(|A|,\bar{A}\right);\ \left[B=\left(|B|,\bar{B}\right)\right].$$

The purpose of the system is the most profitable (V,W) — the exchange; those. for the minimum amount of resource V to obtain the largest possible amount W, which is a function of the structure and behavior of both systems:

$$W=W\left(V,|A|,|B|,\bar{A},\bar{B}\right)=W\left(V,A,B\right).$$

As a result of the interaction between the system A and B receive the following (V,W) — exchanges:

$$\underline{W}_a=W_a\left(V_a,A_0,B_0\right)=$$

$$=\max_{\{\bar{A},|A|\}}\left[\left(1+\alpha\right)W_a\left(V_a,A,B\right)-\alpha\max_{\{\bar{B},|B|\}}W_a\left(V_a,A,B\right)\right];$$

$$\underline{W}_b = W_b\left(V_b, A_0, B_0\right) =$$

$$= \max_{\{\bar{B},|B|\}}\left[\left(1+\alpha\right)W_b\left(V_b, A, B\right) - \alpha \max_{\{\bar{A},|B|\}} W_b\left(V_b, A, B\right)\right];$$

Where

$$\alpha = \begin{cases} -1 \; - \; \textit{в конфликтной ситуации}; \\ 0 \; - \; \textit{в индифферентной ситуации}. \end{cases}$$

Here A_0 and B_0 are optimal systems, i.e., systems whose structures and behavior are optimal, since with their help it is possible to produce (V,W) — exchanges that are close to optimal.

To determine its goal, each system needs to decide what is more important for it, whether to get the most profitable (V,W) **exchange** or to prevent another system from doing it. In this case, the systems can vary the values (V,W) of exchanges within certain limits:

$$\underline{W}_1 \le W_a \le \overline{W}_1;$$

$$\underline{W}_2 \le W_b \le \overline{W}_2,$$

Where $\underline{W}_1$ And $\underline{W}_2$ correspond to the most aggressive systems, and $\overline{W}_1$ and $\overline{W}_2$ to the most cautious ones. If the goals of the systems are known, then there is a well — defined situation. If each system or one of them hides its intentions, then there is a game situation regarding the choice of the target.

Denote by $\underline{A}_i$ and $\underline{B}_j$ $(i=1,...,n; \quad j=1,...,m)$ the goals of the systems A and B respectively. The goals of $\underline{A}_1$ and $\underline{B}_1$ consist in causing maximum damage to the opposite system, and the goals of $\underline{A}_n$ and $\underline{B}_m$ correspond to the extreme caution of both systems. All other goals $\underline{A}_i$ and $\underline{B}_j$ correspond to intermediate situations and are numbered in the order of successive transition from $\underline{A}_1$ to $\underline{A}_n$ and from $\underline{B}_1$ to $\underline{B}_m$. Assuming that in the situation $\{\underline{A}_i, \underline{B}_j\}$ the systems A and B receive a payoff $\underline{W}_a = a_{ij}$ and $\underline{W}_b = b_{ij}$, we obtain a bimatrix game to determine the optimal goal with payoff matrices $\|a_{ij}\|$ and $\|b_{ij}\|$. If we accept that $a_{ij} \le a_{sj}$ for $i \le s$; $b_{ij} \le b_{sj}$ as $i \le s$; $a_{is} \le a_{ij}$ with $j \le s$; $b_{is} \le b_{ij}$ as $j \ge s$, then the solution of the game is trivial and consists in the fact that both systems must adhere to the goals $\underline{A}_n$ and $\underline{B}_m$, i.e., be careful and not be aggressive. The interaction of complex systems becomes strictly antagonistic when (V,W) — exchange is positive for one system and negative for the other, and $\underline{W}_a = \underline{W}_b$.

As a rule, the interaction of systems A and B during (V,W) — exchange is stochastic in nature, and therefore we can only talk about a certain probability $P(V,W)$ of each system achieving its goal. This probability serves as an indicator of the efficiency of the system's behavior. The maximum value of this probability is defined as marginal efficiency.

Systems A and B cannot expect to receive $W_a > \overline{W}_a$ and $W_b > \overline{W}_b$. Therefore, within the limit

$$P_a(V_a, W_a) \cong 0 \text{ at } W_a > \overline{W}_a;$$
$$P_b(V_b, W_b) \cong 0 \text{ at } W_b > \overline{W}_b.$$

Similarly, values $W_a < \overline{W}_a$ and $W_b < \overline{W}_b$ systems can be obtained with high probabilities, i.e. within the limit

$$P_a(V_a, W_a) \cong 1 \text{ at } W_a < \overline{W}_a;$$
$$P_b(V_b, W_b) \cong 1 \text{ at } W_b < \overline{W}_b.$$

Such an asymptotic behavior of the probability $P(V,W)$ makes it possible in a specific situation to find the limiting law (V,W) of — exchange, which establishes the limits of the efficiency of the system's functioning.

All elements A and B are divided into three classes:

— working (vital) a — and b — elements;

— protective R_a — and R_b — elements;

— active (influencing the external environment) c_a — and c_b — elements.

Let's take an example. We will assume that before the interaction of the system A and B have some limited “resource reserves” $\overline{V}_a$ and $\overline{V}_b$, from which vital elements are generated, and

$$\overline{V}_a = \{V_{aj}\}; \overline{V}_b = \{V_{bj}\}; \ \left(j = 1, \dots, n_a\left(n_b\right)\right),$$

where n_a is the number of types of vital elements of the system A; n_b is the number of types of vital elements of the system B.

From $V_{ai}\left(V_{bj}\right)$ can be played $A_i\left(B_j\right)$ $a(b)$ — elements j of the — th type of values $a_j(b_j)$. The values of system elements are determined in relative units, chosen on the basis of expert assessments and in each case according to their own rules.

The protective and active elements of each system are reproduced (generated) by vital elements.

Initial resources $\overline{V}_a$, and $\overline{V}_b$ due to their limitations, can reproduce either many ineffective and low — cost elements (for example, buying advertising links on a large number of low — rated web resources), or few highly effective and high — cost elements (for example, buying advertising links in the prestigious Yandex catalog).

Thus, in the information system there are certain critical relationships between the number of all elements that ensure maximum survivability.

We will assume that the systems A And B at the time of the beginning of the interaction have:

1) A_j and B_j vital elements of type ($j = 1, \dots, n_a(n_b)$) and values a_j and b_j respectively, and

$$\sum_{j=1}^{n_a} a_j A_j = N_a(0); \quad \sum_{j=1}^{n_b} b_j B_j = N_b(0);$$

2) r_a and r_b types of protective elements by α_m and β_m, where

$$\sum_{m=1}^{r_a} \alpha_m = N_{R_a}(0); \quad \sum_{m=1}^{r_b} \beta_m = N_{R_b}(0);$$

3) S_a and S_b types of protective elements according to V_m^a and V_m^b, and

$$\sum_{m=1}^{S_a} V_m^a = M_a(0); \quad \sum_{m=1}^{S_b} V_m^b = M_b(0).$$

Systems interaction A And B occurs over a certain period of time and consists in the mutual exchange of certain portions of active c elements.

Systems A and B at each moment of time t determine their behavior:

$$\bar{A}(t) = \left\{ \mu_\omega^a(t), \sigma_\omega^a(t) \right\};$$

$$\bar{B}(t) = \left\{ \mu_\omega^b(t), \sigma_\omega^b(t) \right\},$$

where $\mu_\omega^{a(b)}(t)$ and are $\sigma_\omega^{a(b)}(t)$ portions of $R_a(R_b)$ — and $c_a(c_b)$ — elements aimed at protecting and/or destroying ω — elements (ω — element – $\left(\omega = a, b, R, C\right)$).

Over time, portions of c_a — and c_b — elements fill the systems B and, A respectively, and, thus, over time, the elements of the system become thinner if there is no replenishment.

To obtain a quantitative assessment of the interaction of systems A and B it is necessary to specify the meaning (V, W) of — exchanges, which both systems strive to optimize. We assume that the survivability of the system is determined by the presence of vital elements in it. In other words, the system $A(B)$ is functioning normally at the moment if the following conditions are met: t

$$Q_a(A_1,\dots,A_{n_a},t) > Q_a \min(t);$$
$$Q_b(B_1,\dots,B_{n_b},t) > Q_b \min(t),$$

where Q_a and Q_b are given functionals of time and structures of systems A and, B respectively.

In particular, one can take

$$Q_a = \sum_{j=1}^{n_a} a_j A_j(t); \quad Q_a \min(t) = \sum_{j=1}^{n_a} Q_{aj} a_j A_j(0);$$
$$Q_b = \sum_{j=1}^{n_b} b_j B_j(t); \quad Q_b \min(t) = \sum_{j=1}^{n_b} Q_{bj} b_j B_j(0),$$

where $0 \le Q_{aj}, Q_{bj} \le 1$ are the parameters determined by the specifics of the systems and the requirements that apply to their survivability.

The choice of optimality criteria in the above form seems natural in those situations where the survivability of the system depends strongly on the total value of vital elements and, to a lesser extent, on the distribution of elements by their types.

Parameters Q_{aj} and Q_{bj} characterize the viability of systems A and, B respectively. Indeed, if the system A is functioning normally $(1-Q_{aj})$ at the moment , then t the — th fraction a of — elements j of the — th type $(j=1,\dots,n)$ is in working condition at that moment. At Q_{aj} = 0 $(j=1,\dots,n)$, the system A is the most viable, and at , $Q_{aj} \to 1$ the viability of the system A tends to zero. Likewise for the system B .

If we take into account that the exchange processes (V,W) of each system are interconnected, then the situation of the interaction of systems A and B can be considered as a generalized conflict, for the description of which it is necessary to set the payoff function. The form of the payoff function of each system essentially depends on the degree of conflict in the situation, i.e. on how one system affects the feasibility of the goal by another system, as well as on the level at which the conflict is fought.

In the general case, taking into account possible changes in goals, the intensity of the conflict can also change the form of the payoff function over time. Thus, the considered generalized conflict situation between systems A and B is a zero — sum game of two players with the following payoff functions:

$$Q_a(T) = \sum_{j=1}^{n_a} a_j A_j(T); \quad Q_b = \sum_{j=1}^{n_b} b_j B_j(T)$$

when fixing the end time of the conflict, and when the optimality criteria are non — autonomous, the game ends at the time $t^* = \min(t_a, t_b)$, where t_a, t_b are the smallest non — negative roots of the equations $Q_a = Q_{a\ \min}$ and $Q_b = Q_{b\ \min}$.

The outcome of the conflict essentially depends on the degree of importance for each system of posing the question: whether to eliminate the enemy or survive by any means, i.e. depending on the intensity of the conflict, the payoffs of both systems can vary within certain limits:

$$\underline{W}_1 \le \max_{(\bar{A},A)} \min_{(\bar{B},B)} \sum_{j=1}^{n_a} a_j A_j(T) \le \bar{W}_1;$$

$$\underline{W}_2 \le \max_{(\bar{B},B)} \min_{(\bar{A},A)} \sum_{j=1}^{n_b} b_j B_j(T) \le \bar{W}_2.$$

Let the interaction of systems A and B occur on a certain time interval $0 \le t \le T$ at discrete moments t_i. Systems A and B at each moment of time t determine their behavior:

$$\bar{A}(t) = \left\{ \mu_\omega^a(t), \sigma_\omega^a(t) \right\};$$

$$\bar{B}(t) = \left\{ \mu_\omega^b(t), \sigma_\omega^b(t) \right\};$$

where $\mu_{\omega}^{a(b)}(t)$ and are $\sigma_{\omega}^{a(b)}(t)$ portions of $R_a(R_b)$ — and $c_a(c_b)$ — elements aimed at protecting and destroying ω — elements. Let's designate a portion of active elements: $\sigma_{\omega}^{a}=\left(\sigma_{1\omega}^{a},...,\sigma_{s_a\omega}^{a}\right)$; $\sigma_{\omega}^{b}=\left(\sigma_{1\omega}^{b},...,\sigma_{s_b\omega}^{b}\right)$ and a portion of protective elements: $\mu_{\omega}^{a}=\left(\mu_{1\omega}^{a},...,\mu_{s_a\omega}^{a}\right)$; $\mu_{\omega}^{b}=\left(\mu_{1\omega}^{b},...,\mu_{s_b\omega}^{b}\right)$.

In these notations, at the time t_i of the strategy of the behavior of the systems A and B will have the form

$$\bar{A}^{(i)}=\left\{\left\|\mu_a^a\right\|,\left\|\mu_c^a\right\|,\left\|\sigma_b^a\right\|,\left\|\sigma_R^a\right\|,\left\|\sigma_c^a\right\|\right\};$$
$$\bar{B}^{(i)}=\left\{\left\|\mu_b^b\right\|,\left\|\mu_c^b\right\|,\left\|\sigma_a^b\right\|,\left\|\sigma_R^b\right\|,\left\|\sigma_c^b\right\|\right\}.$$

The change in the average number of survivors until the moment t_{i+1} of elements will be described by the following relationships:

— vital and protective for A:

$$A_n(t_{i+1})=\max\left\{0,A_n(t_i)-\sigma_{an}^b(t_i)P_{an}^b(t_i)\right\}\quad (n=1,...,n_a);$$
$$\alpha_j(t_{i+1})=\max\left\{0,\alpha_j(t_i)-\sigma_{R_j}^b(t_i)P_{R_j}^b(t_i)\right\}\quad (j=1,...,r_a).$$

— vital and protective for B:

$$B_m(t_{i+1})=\max\left\{0,B_m(t_i)-\sigma_{bm}^a(t_i)P_{bm}^a(t_i)\right\}\quad (m=1,...,m_b);$$
$$\beta_s(t_{i+1})=\max\left\{0,\alpha_s(t_i)-\sigma_{R_s}^a(t_i)P_{R_s}^a(t_i)\right\}\quad (s=1,...,r_b),$$

where the probability of failure of the corresponding active elements

$$P_{an}^b(t_i)=\left(P_{1an}^b,...,P_{s_ban}^b\right);\quad P_{R_j}^b(t_i)=\left(P_{1R_j}^b,...,P_{s_bR_j}^b\right);$$
$$P_{bm}^a(t_i)=\left(P_{1bm}^a,...,P_{s_abm}^a\right);\quad P_{R_s}^a(t_i)=\left(P_{1R_s}^a,...,P_{s_aR_s}^a\right).$$

At the same time, taking into account the possibility of failure c — elements in each system, we obtain the following limiting conditions:

$$\sum_{i=0}^{N-1}\left\{\sigma_{cn}^{b}(t_i)P_{cn}^{b}(t_i)+\sum_{j=1}^{n_b}\sigma_{nb_j}^{a}(t_i)+\sum_{s=1}^{r_b}\sigma_{nR_s}^{a}(t_i)+\sum_{j=1}^{S_b}\sigma_{nc_j}^{a}(t_i)\right\}=v_n^a(0),\quad (n=1,...,S_a);$$

$$\sum_{i=0}^{N-1}\left\{\sigma_{cs}^{a}(t_i)P_{cs}^{a}(t_i)+\sum_{j=1}^{n_a}\sigma_{sa_j}^{b}(t_i)+\sum_{j=1}^{r_a}\sigma_{nR_j}^{b}(t_i)+\sum_{j=1}^{S_a}\sigma_{sc_j}^{b}(t_i)\right\}=v_n^b(0),\quad (s=1,...,S_b),$$

Where $P_{cm}^{b}(t_i)=\left(P_{1cn}^{b},...,P_{s_bcn}^{b}\right);\quad P_{cs}^{a}(t_i)=\left(P_{1cs}^{a},...,P_{s_acs}^{a}\right).$

The above equations characterize the state of systems A and B on average. The solution of these equations allows one to find strategies and determine the conditions for the survival of each system.

When building a model of interaction between two systems A and B it is necessary to take into account such factors as the state of awareness of the parties and the possibility of preventive actions. The information factor plays an important role in choosing the optimality criterion and quite strongly influences the quality of optimal strategies.

Let us assume that the systems A and B evaluate the situation based on the payoff functions Q_a and Q_b. Then the original game can be replaced by an equivalent game with payoff function $Q=Q_a-Q_b$. The system A will seek to maximize Q and the system B will seek to minimize. In this case, the following situations are possible:

1. The system B is informed about the choice of strategy $\bar{A}(t)$ by the system A. The system A does not have such information. Therefore, the optimal strategy $\bar{A}_{opt}(t)$ of the system A must be determined from the condition

$$\max_{\bar{A}(t)}\min_{\bar{B}(t)}Q\left[\bar{A}(t),|A(t)|,\bar{B}(t),|B(t)|\right]=$$

$$=\max_{\bar{A}(t)}Q\left[\bar{A}(t),|A(t)|,\bar{B}*\left[t,\bar{A}(t)\right],|B(t)|\right]=$$

$$=Q\left[\bar{A}_{opt}(t),|A(t)|,\bar{B}*\left[t,\bar{A}_{opt}(t)\right],|B(t)|\right]=V_1\left[|A(t)|,|B(t)|\right].$$

2. The system A is informed about the choice of strategy $\overline{B}(t)$ by the system B. The system B does not have such information. In this case, the optimal strategy $\overline{B}_{opt}(t)$ of the system B can be determined from the condition

$$\min_{\overline{B}(t)} \max_{\overline{A}(t)} Q\left[\overline{A}(t),|A(t)|,\overline{B}(t),|B(t)|\right]=$$
$$=\min_{\overline{B}(t)} Q\left[\overline{A}^*\left[t,\overline{B}(t)\right],|\overline{A}(t)|,\overline{B}(t),|B(t)|\right]=$$
$$=Q\left[\overline{A}^*\left[t,\overline{B}_{opt}(t)\right],|A(t)|,\overline{B}^*\left[t,\overline{A}^*_{opt}(t)\right],|B(t)|\right]=V_2\left[|A(t)|,|B(t)|\right].$$

3. Systems A and B have complete information about the choice of strategy for each of them. Then the optimal strategies $\overline{A}^*_{opt}(t)$ and $\overline{B}^*_{opt}(t)$ systems A and B can be found from the relation

$$\max_{\overline{A}(t)} \min_{\overline{B}(t)} Q\left[\overline{A}(t),|A(t)|,\overline{B}(t),|B(t)|\right]=$$
$$=\min_{\overline{B}(t)} \max_{\overline{A}(t)} Q\left[\overline{A}(t),|A(t)|,\overline{B}(t),|B(t)|\right]=$$
$$=Q\left[\overline{A}^*_{opt}\left[t,\overline{B}^*_{opt}(t)\right],|A(t)|,\overline{B}^*\left[t,\overline{A}^*_{opt}(t)\right],|B(t)|\right]=$$
$$=V_3\left[|A(t)|,|B(t)|\right].$$

It follows from the above

$$V_1\left[|A(t)|,|B(t)|\right]\le V_3\left[|A(t)|,|B(t)|\right]\le V_2\left[|A(t)|,|B(t)|\right].$$

Thus , the value of mutual tactics of awareness of systems A and B is determined by the difference $\Delta V = V_2 - V_1$.

2.4. Logical — probabilistic models

To analyze and assess the survivability of information systems that operate in NI conditions, one can use logical — probabilistic models, according to which it is assumed that the elements of the system and the system itself have a two — valued logic of operation, and all events in the system are independent. The description of the system is possible with the help of a static model that does not contain the time parameter among the independent variables. Functional dependencies between variables can be fully represented using the functions of the algebra of logic. The results of the action of influences are also evaluated using a binary scheme: either the state from the set of operable states is preserved, or the operability is disrupted (the system goes into a state from the set of inoperable states).

Let us assume that the elements of the system are point objects, which are interconnected by communication lines. The sequence of NV of impulse type forms a stream of independent events. There are no secondary consequences of NI, so the established state of the system is known immediately after NI.

Consider a system that consists of K_s objects s — but —

measures of a variant of the system in which, after a one — time unwanted impact on i the — th object damage occurs c_i^s .

Objects are numbered for each variant of the system as the damage decreases, i.e. $c_1^s > c_2^s > ... > c_n^s$. Set the threshold allowable damage A and we will assume that in the case of multiple NV, different objects can be influenced and, first of all, objects, the influence on which brings the greatest damage.

The damage to the system as a whole is the sum of damages from influences on individual objects. The system damage index K_s is calculated as follows:

$$K_s^A = \min_{C_s \geq A} K_s; \quad c_s = \sum_{i=1}^{K_s} c_i^s .$$

If K_s^A it exceeds the threshold allowable damage value A, then the system becomes inoperable.

Let some information system having a basic structure S_0 performs a task for some time t. After NV, the structure of the system may change, i.e. the system will have a new structure S_i with which it can be operable, in other words, the structure belongs to the set of operable structures:

$$S^P = \left\{ S_i, i = 1, \ldots, N_p \right\}.$$

Or the system may be inoperable, in which case S_i it belongs to a set of inoperable structures, namely

$$S^{NP} = \left\{ S_i, i = N_{p+1}, \ldots, N \right\}.$$

Thus, after an NI, either a workable procedure from the set S^P or an inoperable procedure from S^{NP}.

After n the — fold HB, the system with the appropriate structure should start executing the task, and its execution time should be t.

Let's assume that each element of the system can be in one of three states: e_0 — the element is operational and put into operation; e_1 — the element is operational, but not included in the work for some reason; e_2 — the element is inoperative. The connections between the elements are defined and stationary in time, i.e. at an arbitrary point in time, the state of an element can be determined by the health state of this element and the states of other elements. Signs of system performance are unchanged in time and allow you to simultaneously determine the state of the system by the totality of the states of its elements.

There are such main stages of the system survivability analysis based on a logical — probabilistic model.

Description of element states. For each element, two logical variables are introduced: x_i — health indicator i of the — th element:

$$x_i = \begin{cases} 1, & i\text{-é ýëåìåíò ðàáîòîñïîñáåí}; \\ 0, & i\text{-é ýëåìåíò íåðàáîòîñïîñîáåí}; \end{cases}$$

y_i — status indicator i of the — th operable element:

$$y_i = \begin{cases} 1, & i\text{-é ýëåìåíò ðàáîòàåò}; \\ 0, & i\text{-é ýëåìåíò íå ðàáîòàåò}. \end{cases}$$

To display the action of influence on the elements, indicators z_{ij} and are also introduced $z_i = \bigcup\limits_{(j)} z_{ij}$, where

$$z_{ij} = \begin{cases} 1, & \text{êî ãäà âëèÿí èå } j\text{-ãî ò èï à} \\ & \text{äåéñò âóåò í à } i\text{-é ýëåì åí ò;} \\ 0, & \text{êî ãäà âëèÿí èå } j\text{-ãî ò èï à} \\ & \text{í å äåéñò âóåò í à } i\text{-é ýëåì åí ò.} \end{cases}$$

Next, indicators of the three states of the element are defined, which are then used to compose logical dependencies:

$$u_{i0} = x_i y_i \bar{z}_i;$$
$$u_{i1} = x_i \bar{y}_i \bar{z}_i;$$
$$u_{i2} = \bar{x}_i \vee x_i z_i.$$

Compilation of logical dependencies. Based on a preliminary analysis of dynamic models of information processes, taking into account the counteraction of NV, reconfiguration and control, a system of logical equations can be compiled, which can be represented in vector form:

$$Y = f_Y\left(X, Y, Z\right).$$

For an information system, the state of a relatively small group of the most significant initial elements is decisive, however, due to the presence of indirect links, the system's performance is also determined by the state of all other elements.

The system health function (FSF) is written as

$$F = \mathop{\wedge}\limits_{(i)} f_i\left(X, Y, Z\right),$$

where f_i — logical function ($i = 1,.., N$) — progress indicator of i the th function of the system.

Solving a system of logical equations. The system considered above is linear, which can be solved, in particular, by methods of determinants, substitutions, etc. [33].

Probabilistic description of elements and external influences. Each element in a probabilistic model is the probability $p_i = P(x_i = 1)$ that, at a given arbitrary time, the element x_i workable. With the advent of influences $z_{ij} = 1$.

The resistance of the i — th element to j the — th excitation can be taken into account using the probability that the element will remain operational in the event of an excitation. In addition, separate in the probability of hitting the element x_i in the area of action of *the* j — th NV factor.

Transforming the Fed to form a transition to substitution. There are forms of transition to partial and complete replacement. Complete substitution forms: complete disjunctive normal form, disjunction of orthogonal non — repetitive forms, non — repetitive form in the "AND — NET" basis, etc. [33] After reduction to one of these forms, a one — step substitution of logical variables and operations for probabilities and arithmetic operations is performed.

Mixed form entry. The replacement of non — repetitive variables in the transformed FRS is, as a rule, not a complete, but a partial replacement, as a result of which some logical variable and operation are replaced by a probability and an arithmetic operation. The FRS obtained in this way is called the mixed form, since it contains both logical variables and probabilities.

Determination of indicators of survivability. With the help of a multi — step procedure for replacing logical variables in mixed forms, which are compiled for the base S_0 and structures S_i, we find the probabilities of operable states at the time t — $P(t/S_0)$ and $P(t/S_i)$, as well as the conditional function of survivability $G_i(t)$:

$$G_i(t) = \frac{P(t/S_i)}{P(t/S_0)}.$$

After that, one can find the survival function, the unconditional survivability function, and the amount of HB (average) that brings the system to an inoperable state.

Having defined the survivability function as a function of system survival after n a — fold NI (event A_n), averaged over all possible structures, $P_n(k)$ is the average probability of occurrence of the structure S_n after n — fold HB, one can write

$$G\left(t/A_n\right) = G\left(t,n\right) = \sum_{k=1}^{N} P_n(k) G_k(t).$$

The unconditional survival function is an average over all possible events A_n system survival features:

$$G(t) = \sum_{n=1}^{\infty} P(A_n) G\left(t/A_n\right) = \sum_{k=1}^{N} P(S_k) G_k(t)\ ,$$

where is the probability $P(S_k)$ is determined by the formula

$$P(S_k) = \sum_{n=1}^{\infty} P(A_n) P_n(k).$$

After that, one can find the survival function, the unconditional survivability function, and the amount of HB (average) that brings the system to an inoperable state.

Next, we determine the system survival index after n the — fold NI (in the event of the occurrence of an event A_n):

$$R(n) = 1 - Q(n) = P(F = 1/A_n),$$

Where $Q(n)$ — conditional law of vulnerability: $Q(n) = P(F = 0/A_n)$.

The average number of NV, leading to the loss of system performance, is found by the formula

$$\bar{\omega} = \sum_{n=1}^{\infty} R(n),$$

and the average survivability, $\bar{d} = \bar{\omega} - 1.$

In turn, the average number of NIs leading to failure to complete the task can be determined as follows:

$$\bar{\omega}(t) = \sum_{n=1}^{\infty} R(n) = \sum_{n=1}^{\infty} n\big(G(t, n-1) - G(t, n)\big) = \sum_{n=1}^{\infty} G(t, n).$$

2. 5. Estimation of the survivability of the system according to its state

Consider an information system that consists of N elements with arbitrary connections between themselves and the health function $F_i = f(X),\quad X = \{X_1, X_2, ..., X_N\}$. The system is under the action of a flow of independent point NV:

$$\varphi_{kj} = \frac{1}{N},$$

Where φ_{kj} is the probability that k the — th element will fall into the scope of j the — th NI, i.e. it is assumed that each element is equally likely to be vulnerable to NI.

Let us estimate the survivability of such a system. System survival function $R(n)$ after n — fold HB can be represented as

$$R(n) = \sum_{X \in X'} P(x / A_n) = P(F = 1 / A_n),$$

Where X' is a set of vectors X that correspond to the operational states of the system. Probability $P(x / A_n)$ is found according to the formula

$$P(x/A_n) = \sum_{\vec{n}\in M_n} P(\vec{n})P(X/\vec{n}),$$

where $\vec{n} = (n_1, n_2, ..., n_k)$ is the vector of the number of NIs that fall on k the subsystems; M_n is the set of vectors satisfying the condition $n_1 + n_2 + ... + n_k = n$. Probability $P(\vec{n})$ calculated by the formula

$$P(\vec{n}) = \frac{n!}{n_1!n_2!...n_k!}\gamma_1^{n_1}\gamma_2^{n_2}\cdots\gamma_k^{n_k},$$

where γ_i **is** the probability that i the th subsystem is included in the scope of the NV.

Given the uniformity of the vulnerability of the elements of the system, the above expressions can be simplified. We represent the FRS in the form of an orthogonal disjunctive normal form: $F = \overset{m}{\underset{i=1}{\vee}} Q_i$.

Respectively,

$$R(n) = \sum_{i=1}^{m} P(Q_i = 1/A_n).$$

For the case of equiprobable entry of elements into the scope of the HB, based on the basic structure S_0, all possible workable structures are determined $S^p = \{S_i, i = 1, 2, ..., N_p\}$, and the survival function takes the form

$$R(n) = \sum_{j=1}^{N_p} \frac{r_j(n)}{N^n},$$

where $r_j(n)$ is the number of cases in which a structure appears S_j after n a — fold NI; $r_j(n)$ can be determined by the formula

$$r_j(n) = \sum_k L_{nk}B_{kj},$$

where L_{nk} is the number of permutations of n type B_{kj} elements ; k is the number of different vectors X with k zeros that lead to the structure S_j .

Numbers L_{nk} can be calculated directly from the formula

$$L_{nk} = \sum_{i=1}^{k} C_k^i i^n (-1)^{k+1}.$$

In the Boolean model, it can be assumed that the infoplot consists of n elements (documents), while i the — th element corresponds to the bu — left variable x_i , which can take the values {0, 1}, i.e.

$$x_i = \begin{cases} 1, \text{ если элемент } i \text{ активен,} \\ 0, \text{ иначе.} \end{cases}$$

Taking into account the fact that the system is dynamic, we can fix the value of n as a deliberately large number that exceeds the maximum observed number of documents in information plots. Non — existent (missing) documents can be assigned zero values x_i .

The state of the information system is determined by the structural Boolean function of its performance (effectiveness), depending on the variables $x_1, x_2, \ldots, x_n$:

$$S\left(x_1, x_2, ..., x_n\right) = \begin{cases} 1, \text{ если система активна,} \\ 0, \text{ иначе.} \end{cases}$$

If the activity of an element of an information system is considered as a function of time, then the state of *the i* — th document can be considered as a random process $x_i(t)$ that takes the values 0 and 1 at arbitrary times. $t \geq 0$ For an information system, the probability of its performance is determined according to known rules [34, 35].

Among the shortcomings of Boolean models, one can name the assumption of only two states of the components – active and inactive. In addition, Boolean models do not take into account the fact that the sequence in which individual components fail can play a very significant role. In addition, in the general case, the nature of failures of individual components depends on the state of other components. This is in conflict with the initially assumed independence of the elements in the Boolean model.

An information system can also be described by a Markov model. Let the system have m possible states. Let us denote the set of states as $M=\{z_1,\ z_2,...,z_m\}$. For any fixed moment of time, $t\geq 0$ the state of the system $z_i(t)$ is interpreted as a random variable. The set of all states M, the initial probability distribution vector p (0), and the transition probability function are given. The probability of the relevance of the information system at a given time t (system performance) is determined [34].

The applicability of Markov models also has its limits. The intensity of transitions between individual states of the system can be non — stationary, the assumptions made in the calculation regarding the distribution of the failure rate can significantly reduce the accuracy of the results obtained; the number of states of the system can be so large that the calculation becomes practically impossible.

After evaluating the reliability of the system components and obtaining general indicators of its reliability, it is possible to assess its survivability at all stages of their life cycle. There are several general approaches to survivability assessment . The survivability of the system can be assessed with respect to some standard external influence or with respect to a set of external influences [36].

Let be $E=\{e_i\}$ the set of destructive influences on the information system; $\sigma_j(e_i)$ is an indicator of the effectiveness (quality) of the functioning j of the th version of the information system under the influence of e_i; $H_j(E)=\min\limits_{e_i\in E}\sigma_j(e_i)$ is an indicator of the survivability of the information system for a variety of possible impacts on it E. Then, with the purposeful formation of an information system, the design task is to Ω find from a variety of options for information systems one for which

$$H_k(E)=\max_{X_j\in\Omega}\min_{e_i\in E}\sigma_j(e_i).$$

In addition to the need to preserve many functions of information systems under adverse impacts on the information system, the task of maintaining a certain level of its effectiveness (relevance, informativeness) is often set.

There are numerous approaches to quantitative assessment of survivability , the most common of which is to determine the ratio of the number of functional (workable) states of the system to the total possible number of system states arising under destructive impacts.

As a simple example of an information system, consider an information plot consisting of four documents ($n = 4$). Destructive impact on the information plot is the elimination of the documents included in it from the information space. Moreover, the first of the documents is considered decisive – its removal from the information space actually leads to the loss of the information functionality of the entire information plot. The remaining three documents are considered equal. Elimination of any two of them ($k = 4$) also leads to the loss of functionality of the information plot.

If we denote the state of the information plot by a 4 — element tuple, then the set of inoperative states can be represented as the union of two subsets, the first of which corresponds to the states with the first document eliminated, and the second – with the actual first document, but missing the other two.

The power of the first subset is $2^{n-1} = 8$, we list its components:

(0, 0, 0, 0)
(0, 0, 1, 0)
(0, 0, 0, 1)
(0, 0, 1, 1)
(0, 1, 0, 0)
(0, 1, 1, 0)
(0, 1, 0, 1)
(0, 1, 1, 1).

The power of the second subset is $C_n^k + 1 = 4$, its components are:

(1, 0, 0, 0)
(1, 0, 1, 0)
(1, 0, 0, 1)
(1, 1, 0, 0).

The power of the entire set of states after the destructive impact is 2^n $^-1 = 15$.

Thus, the survivability G of the infoplot is

G \u003d (15 – 8 – 4) / 15 \u003d 0.2.

Let us consider the case when all states of the infoplot after the destructive impact are equivalent, i.e. equiprobable.

If the states of the information plot are not equiprobable, then the survivability G of the information plot is

$$G = \frac{\sum_{i=1}^{m} p_i - \sum_{i=1}^{j} p_i^{(0)} - \sum_{i=1}^{l} p_i^{(1)}}{\sum_{i=1}^{m} p_i} = 1 - \sum_{i=1}^{j} p_i^{(0)} - \sum_{i=1}^{l} p_i^{(1)},$$

where m is the power of the entire set of states after the destructive impact ($m = 2^{n-1}$); j is the power of the subset of states with the first plot eliminated; l is the power of the subset of states with the actual first plot ($l = C_n^k$ + 1); p_i is the probability *of the i* — th state after the destructive impact; $p_i^{(0)}$ — probability of *the i* — th

states after the destructive impact and elimination of the first plot; $p_i^{(1)}$ — the probability of *the i* — th state after the destructive impact and the preservation of the first plot, but with the loss of the general relevance of the infoplot.

The presented model allows one to operate with numerical results, but is too simplified. Turning to more realistic approaches, it should be noted that when characterizing an information system, it is necessary to pay attention to such properties as the divisibility or weakening of an information message, flow, impact on a heterogeneous information space, non — stationarity of this impact, dynamic changes in the areas of the information space itself.

2.6. Survivability rating according to the results of the task

For the information system to perform a common task and achieve the system goal, each element must function in accordance with the established schedule, ensuring the performance of a given set of functions over time t. Let the probability that the i — th element performs the functions defined for it is $p_i(t)$. In the general case, $p_i(t)$ there may be a probability of performing a task with a complex mode of operation and complex functioning restrictions.

Assessing the survivability, we will assume that all NV have ended by the beginning of the time interval $(0,t)$.

Vitality is assessed as follows. Based on the logical — probabilistic model, the FRS is written for the basic structure. Then all other workable structures of the system are identified, substituting vectors into the FRS in which one, then two, three and more elements are replaced by zeros. If, after substitution, the logical function does not become identically equal to zero, then it corresponds to one of the workable structures. At the same time, during such tests, the coefficients are determined B_{ki}. multiplying B_{ki} onto the matrix $\|L_{nk}\|$, we find the matrix of coefficients $\|r_{ni}\|$.

The ratio r_{ni}/N^n gives the probability $P_n(i)$ the fact that after n — multiple NV there is a structure $S_i \in S^p = \{S_k, k=\overline{1,N_p}\}$. For each of the structures S_i the health function is built $F^{(i)}(X)$, which is presented in the following form:

$$F^{(i)}(X) = x_n\left(\bigcup_{j=1}^{l} x_j^{f_j^{(i)}(X)}\right),$$

where x_n is a variable that corresponds to the pole of the system; x_j — non — recurring variables; $f_j^i(X)$ are functions of the logic algebra of an arbitrary form.

Next, we turn to the calculation of probabilities:

$$P\left(F^{(i)}(X)=1\right) = p_n\left(1-\prod_{j=1}^{l} q_j^{f_j^{(i)}(X)}\right),$$

Where $p_n = p(x_n = 1),\;\; q_j = 1 - p_j = p(x_j = 0).$

After that, the remaining logical variables are sequentially replaced and the function is obtained $P(t/S_i)$. Then $G_i(t)$ and is determined

$G(t,n)$. If it is possible to determine the probability $P(A_n)$, then we can find the unconditional vitality function.

2.7. Survivability rating for the purpose of functioning

A more general assessment of the survivability of an information system can be built based on the purpose of its operation, a set of informing tasks $Q = \{q_1, q_2, ..., q_m\}$ and a set of components (documents). Indeed, any task $q_i \in Q, \; i = \overline{1,m}$ is characterized by a set of elementary functions (informing about individual aspects) $F_i = \left\{ f_{j_1}, f_{j_2}, ..., f_{j_k} \right\}, \; 1 \le j_k \le n$, from which solutions to this task are built.

Denote by $F = \bigcup_{i=1}^{m} F_i$ the set of sets of elementary functions of the information system. For each informing task, q_i a characteristic of the effectiveness of the solution is given. Let us introduce the function of potentialities of functional modules $\varphi : \{1, 2, ..., p\} \to P(F)$, where $P(F)$ is the set of all subsets F .

To characterize the possible configurations of the information system, we introduce a matrix of potential system capabilities:

$$a_{ij} = \begin{cases} 1, \text{ если } f_i \in \varphi(j), \\ 0, \text{ если } f_i \notin \varphi(j), \; j = \overline{1,p}, \; i = \overline{1,n}. \end{cases}$$

The current configuration of the information system will be characterized by what information functions each module is aimed at. Let us introduce a binary matrix *B in* dimension $n \times p$ — a matrix of the current configuration of the system, such that

$$b_{ij} = \begin{cases} 1, \text{ если модуль } S_j \text{ выполняет функцию } f_i, \\ 0, \text{ в противном случае.} \end{cases}$$

Let us define the module efficiency function $\varphi_{ýô} : I_s \times I_f \times \ \times B \rightarrow T$, where $I_s = \{1,\ 2,...,p\}$ is the set of module indices; $I_j = \{1,\ 2,...,n\}$ is the set of indices of elementary functions; B — set of configuration matrices; T is a numerical set of quantitative measures of efficiency (for example, the size of the audience reading newsplot documents, etc.) If , $\varphi_{эф}(i, j, B) = t_{ij}$ then in the configuration defined by matrix B , the module S_i performs the function f_j with efficiency t_{ij} , $\varphi_{эф}(i, j, B) = 0$, if the module S_i does not perform the function f_j .

To characterize an information system, we introduce the concept of a characteristic state vector — *an n* — dimensional vector (*n* is the power of the set of elementary functions of the system). The initial configuration of the information system, provided that the entire set of functions *F is performed* , will correspond to the characteristic state vector $(0,\ 0,...,0)$. Some current configuration of the information system will correspond to the characteristic vector $(d_1\ , d_2,..., d_i,..., d_n)$, where d_i is the number of "failures" of the function $f_i \in F$. The "failure" of the function $f_i \in F$ is understood as the impossibility of performing the informing function f_i , i.e. d_i — the number of reconfigurations of the information system due to the "failure" of the informing function $f_i \in F$.

Let's solve the problem of finding a set of characteristic state vectors of the information system in which the configuration is implemented that ensures the fulfillment of the purpose of functioning. The power of this set can also serve as a measure of the survivability of the system .

The problem can be solved in two stages.

1. Finding the set of characteristic state vectors of the information system S_f that determine the states in which it is possible to choose a configuration that ensures the execution of a set of elementary functions F . Let some initial information system be characterized by a matrix B_0. .

$$\sum_{j=1}^{p} b_{ij} = 1, \forall i = \overline{1, n},$$

$$\sum_{i=1}^{n} b_{ij} = 1, \forall j = \overline{1, p}.$$

As an optimization criterion, it is natural to choose

$$\Phi = \sum_{i=1}^{n} \sum_{j=1}^{p} b_{ij} \varphi_{эф} \left(i, j, B_0\right) \to \min(\max) .$$

Depending on the specific meaning of the efficiency function, the problem of finding either the maximum Φ or its minimum is valid.

The above expressions describe a problem of a combinatorial type, which can be solved, for example, by the Hungarian method or by using a heuristic algorithm.

Let us assume that there are “failures” of functions, i.e. the state of the information object changes. The new state, taking into account the failures that have taken place, corresponds to the characteristic state vector $\left(d_1,\ d_2,...,d_i,...,d_n\right)$. Recovery is carried out by redistributing functions between its modules. The problem of finding a new system configuration can be described as follows:

$$\Phi = \sum_{i=1}^{n} \sum_{j=1}^{p} b_{ij} \varphi_{эф} \left(i, j, B\right) = \max\left(\min\right),$$

$$\sum_{i=1}^{n} b_{ij} = 1, \forall j = \overline{1, p},$$

$$\sum_{j=1}^{p} b_{ij} = 1, \forall i = \overline{1, n},$$

$$\Phi \ge \Phi^* \left(\Phi \le \Phi^*\right).$$

where Φ^* is the value that determines the minimum (maximum) allowable efficiency.

The desired set S_f includes only the characteristic vectors for which the given problem is solvable. At this stage of the decision, from the set of characteristic vectors of the state of the information system, S_f we select a subset S_q that determines the states of the system in which it is possible to choose configurations that ensure the fulfillment of the goal of functioning.

As an estimate of the survivability of an information system, we can take the cardinality of the set S_q. In the case of information systems, the problem of informing about their various aspects, regardless of the presence or absence of adverse factors, comes first. In this regard, as a quantitative criterion for assessing survivability, it is advisable to use the ratio of the number of functions performed by the system in the presence of certain adverse effects or a multitude of such effects to the total number of system functions, taking into account the criticality of the performed and non — performed functions. The criticality of each specific function is determined individually for each specific information system, based on its specifics. The quantitative indicator of the survivability of a particular information system under given conditions can be calculated by the formula: $S = \sum_{i \in \Delta} \alpha_i / \sum_{j \in \Theta} \alpha_j$, where Θ is the set of all informing functions; Δ — a set of information system functions performed under given conditions ($\Delta \subseteq \Theta$); α_n is the criticality *of the nth* function. Thus, the quantitative assessment of the survivability of the information system will change in the interval [0, 1]. Vitality is the higher, the greater its quantitative assessment.

2.8. Entropy approach to survivability estimation

An information system can be in a variety of acceptable states for it. The system is efficient as long as its structure or organization allows to avoid destructive influences or to localize them, i.e. the survivability of a system depends on its behavior in the state space.

Let the set $\{X\langle m\rangle\}$ form the space of admissible states of the information system. In general $X\langle m\rangle$ can be a random function depending on time: $X\langle m\rangle = X\langle m\rangle(t) == \{x_1(t),..., x_m(t)\}^T$.

Denote by $F\lambda$ the operator of destructive action with the parameter λ. If a destructive impact affects an information system that is in a state of $X^j\langle m\rangle(t) \in \quad \in \{X\langle m\rangle(t)\}_n$, $I = 1,...,n$ and, for example, $F\lambda\left(X^j\langle m\rangle(t)\right) \in \{X\langle m\rangle(t)\}_n$, then we can say that the system "survived" after the impact with the parameter λ. If the information system is organized in such a way that $\forall\lambda: \ F\lambda\left(X^j\langle m\rangle(t)\right) \in \{X\langle m\rangle(t)\}_n$, then the level of system survivability is very high. Thus, the system has the greater level of survivability, the greater the power of the set of admissible states (diversity) [37].

These requirements correspond to W. Ashby's law, according to which the variety of admissible states reduces the variety of inadmissible ones. In addition to the fact that the level of survivability of the system also depends on its desire to stay in the set of admissible states [38], i.e. from its stability.

Any object or system can be considered as a set with diversity. A change in this diversity corresponds to a change in the states of the system, i.e., the composition of its elements, structure or behavior. Thus, the set of system states can be associated with the set of probabilities of these states that has equivalent properties.

The measure of the diversity of system states is directly related to survivability. The entropy of the state of a system is currently widely used as such a measure of diversity [39–41]. It is known that entropy is a measure of the uncertainty of the state of the system, a measure of the lack of information about its actual structure and behavior. In the general case, it is known that the recognition of a system has the least difficulty if the entropy of the state of this system is minimal. In addition, to maximize the complexity of object recognition against the background of the environment, it is necessary to maximize the entropy of the state of the object to the level of the entropy of the environment.

If we use entropy as a quantity characterizing the level of system survivability, then, taking into account the general properties of the entropy function, it has the following properties:

— the entropy characteristic makes it possible to take into account the uncertainty (in the structure and behavior) of the states of the system and, at the same time, is expressed through the probabilistic characteristics of the system;

— the entropy characteristic of the system depends on the dimension of the state space of the system, the number and variety of system elements;

— the entropy characteristic does not depend on the choice of the origin of coordinates in the state space of the system.

In some cases, it is advisable to measure survivability as a relative value:

$$G = \frac{H}{H_{\max}},$$

where $H_{\max}$ is the maximum possible entropy of the system (hereinafter referred to as the record), which characterizes the limiting state of disorder.

Thus, the given indicator of survivability is the ratio of the entropy of the state of the system to the maximum possible entropy of this system under the conditions of restrictions imposed on it.

For all the classes of systems considered so far, $H_{\max}$ there is a single one, although theoretical work is underway to search for such systems for which $H_{\max}$.

Obviously, the determination of the indicator of the level of survivability should be carried out according to the following scheme:

— the method of estimating the value of entropy *H is determined* depending on the type of system;

— the value *of H* of the analyzed system is calculated;

— the restrictions imposed on the system are analyzed;

— within the framework of these restrictions, a system is synthesized with parameters that provide $H_{\max}$;

— the value of the indicator *G is determined* .

Therefore, before evaluating the value of the survivability index, it is necessary to solve the problem of synthesizing this system according to the criterion of maximum uncertainty within the existing restrictions. In some cases, the implementation of the principle of maximum uncertainty provides the lowest costs compared to other methods of action and thereby increases survivability.

The use of the survivability index implies the possibility of assessing survivability, i.e. knowledge of the influence of all components of the system on the very indicator of survivability and their mutual influence on each other in the process of functioning and evolution of the system.

It is true that the level of survivability of the system G remains constant if and only if the relative increase in entropy is equal to the relative increase in the record. It was shown in [42] that during the evolution of the system, the level of its survivability does not change if $H = H_{\max}$. It is also shown there that during the reconstruction or development of the system, its survivability will increase if and only if the relative increment of the entropy of the system is greater than the relative increment of the record. These two statements are combined into one: the survivability of the system does not decrease during the evolution of the system, if the relative increment of the entropy of the system is not less than the relative increment of the system record:

$$\frac{H'}{H} \geq \frac{H'_{\max}}{H_{\max}}.$$

Thus, for a system that has a tendency to increase the level of survivability, its entropy in the process of evolution always tends to the maximum possible.

Let us consider more specific properties of the survivability index of various systems.

Suppose the system is organized so that H = const Always. Therefore, for the above inequality to hold, it is necessary that the condition $H'_{\max} < 0$.

Thus, for H = const for a system seeking to increase its survivability in the process of evolution, resources should be used to reduce the record, its approach to H .

H = 0 for some system. The set of possible states of the system $\{X_{<m>}\}_n = \{x_{<1>},...,x_{<m>}\}_n^T$ and a set of invalid $\{X_{<m>}\}_k = \{x_{<1>},...,x_{<m>}\}_k^T$, together make up a complete group of events; As a result, it is possible, as already noted, to compare the probable measures $P\left(\{X_{<m>}\}_n\right)$ and $P\left(\{X_{<m>}\}_k\right)$.

The system record can be defined as $H_{\max} = -P\left(\left\{X_{<m>}\right\}_n\right) \times \times \ln P\left(\left\{X_{<m>}\right\}_n\right) - P\left(\left\{X_{<m>}\right\}_k\right) \cdot \ln P\left(\left\{X_{<m>}\right\}_k\right)$.

It is necessary to satisfy the condition $H'_{\max} < 0$, which will lead to the minimization of $H_{\max}$, since by assumption $H = 0$. It follows from the basic properties of entropy [38] that $H_{\max} = 0$ if and only if all but one of the probabilities are equal to zero, and this one is equal to I .

Thus, for H = const and $H_{\max} > 0$ the survivability of the system will increase with time if, during the evolution of the system, the probabilities of all system states are redistributed, in which only one state, and with a minimum probability, corresponds to the state of the system, identified with the loss of working capacity.

2. Let $H_{\max}$ = const. To fulfill the above inequality, it is necessary that $H' > 0$, i.e. _ at $H_{\max}$ = const for a system that seeks to increase its survivability in the process of evolution, the entropy of opium must increase and strive for a record.

3. If the evolution of the system proceeds in such a way that $H' > 0$ and $H'_{\max} > 0$ it is easy to see that this case is a combination of the previous two.

4. If $H' > 0$, since $\lim H = H_{\max}$, whose vitality level does not decrease in the course of evolution, the condition

$$\frac{H'}{H'_{\max}} > 1.$$

Thus, the entropy of the system in this case must grow faster than the maximum possible entropy.

5. If the system is such that $H' < 0$, then it is necessary that the condition $H'_{\max} < 0$ be satisfied, which also leads to the general condition:

$$\frac{H'}{H'_{\max}} > 1.$$

Thus, in this case, the entropy of the system must decrease more slowly than the record. This requirement corresponds to the condition for ensuring the entropy stability of dynamic systems introduced by V.L. Strotanovich.

Consider a system of differential equations describing the dynamics of the system:

$$\frac{dx_j}{dt} = f_j\left(x_1,...,x_m,t\right), \quad j=1,...,m.$$

Let's assume the functions are $f_j\left(x_1,...,x_m,t\right)-$ are continuous in some open domain and are differentiable functions of their arguments. Under random initial conditions, a particular solution of the reduced system will be random

time functions. In accordance with [39, 42], the reduced system has a general monotonic entropy stability if, for an arbitrary initial distribution of coordinates, the total entropy of the system decreases monotonically with time. It is known that a necessary and sufficient condition for the general monotone entropy stability of such a system is the fulfillment of the conditions $\frac{dH}{dt} = \sum_{j=1}^{m} M\left[\frac{\partial f_j(x_1,...,x_m,t)}{\partial x_j}\right] < 0$, $dt = x$ for an arbitrary initial distribution law. The resulting condition can be considered in relation to the assessment of the survivability of the system described by the above differential equations.

The proposed formal approach confirmed the fact that a system that reacts to the impact of destructive factors according to a predetermined prescription is hypersensitive to the slightest deviations of the operating conditions from those envisaged and, therefore, cannot have a high level of survivability.

It is determined that the level of system survivability under the influence of destructive factors depends on the power of the set of admissible states of the system and the diversity of this set, and generally corresponds to the law of the necessary diversity of W. Ashby. Proceeding from this, it is permissible to use the ratio of the entropy of the state of the system and the maximum possible entropy under the restrictions imposed on the system as an indicator of the level of survivability. At the same time, the survivability, which is determined by this indicator, does not decrease during the evolution of the system, if the relative increase in the entropy of the system is not less than the relative increase in the record.

In addition, in [37], a close relationship between entropy and traditional indicators of system stability was determined. To use the proposed approach and estimate the level of survivability, it is necessary to estimate the entropy of the analyzed system and, within the constraints existing on this system, synthesize a system with a record and evaluate the latter.

3. COMPLEX INFORMATION NETWORKS

Information systems can be represented as network structures, the so — called dynamic networks [43]. The current state of the information system can be represented as a graph $< M, L >$, where M is the set of components (for example, documents) of the information system, and L is the set of edges, for example, links of similarity, citation, links, etc. The survivability property is directly related to such properties of graphs as connectivity, clustering, the average shortest path between vertices, etc.

At present, along with the traditional theories of graphs, systems, and queuing networks, the theory of complex networks (eng. – *Complex Networks)*[44] is actively developing, within which approaches to solving computationally complex problems characteristic of modern networks are proposed.

The main reason for the relevance of the theory of complex networks is the results of modern work on the description of real computer, biological and social networks. Such networks have characteristics that are not characteristic of networks with equiprobable connectivity of nodes, but are built on the basis of connected structures, power — law distributions, and hub nodes.

The networks of interest are most often sparse—there are only a small fraction of possible edges connecting individual nodes. Therefore, methods of working with sparse matrices are of particular relevance today.

Indeed, almost all modern networks can be considered complex. For example, the well — known problem of synthesizing the topology of a network admits a combinatorial approach based on the representation of the network in the form of a finite graph without loops and multiple edges, whose vertices correspond to network nodes, and the edges correspond to communication lines.

At the same time, the use of graph enumeration methods for solving the problem of topological optimization is considered unpromising, since it is necessary to explore a huge number of possible options for connecting nodes with communication lines. For example, in a network of 10 nodes, there are 2^{45} options for placing communication lines (for 10 nodes, $C_{10}^{2} = \frac{10 \cdot 9}{2} = 45$ connection lines are theoretically possible. Each of these possible communication lines can actually exist – state "1", or not exist – state "0", i.e. all possibilities 2^{45}).

For a smaller number of nodes (for example, $n = 3$) communication lines can actually be counted as ($2^{\frac{3 \cdot 2}{2}} = 8$) variants (Fig. 9).

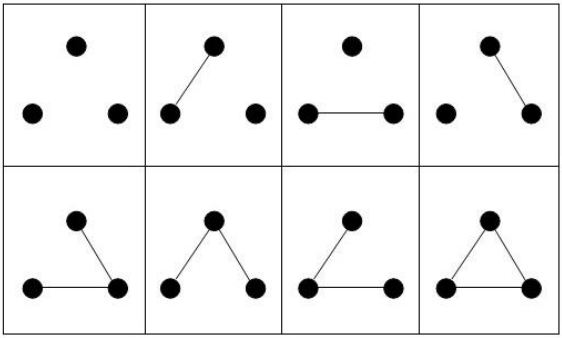


Figure 9. Options for placing communication lines with $n = 3$

Complex networks are usually considered in an abstract space in which the location of the vertices does not matter. For some real types of networks, such consideration is justified.

However, there are many systems in which the location of the components is very important as it affects the evolution of the network. Such networks are called geographical or spatial. In geographic networks, the existence of a direct connection between vertices may depend on many constraints, such as distance between them, geographic topography, territorial restrictions, and so on. Models intended to represent such networks must take these limitations into account.

3.1. Parameters of complex networks and problems of survivability

The theory of complex networks, as a field of discrete mathematics, studies the characteristics of networks, taking into account not only their topology, but also statistical phenomena, the distribution of weights of individual nodes and edges, the effects of leakage, leakage, conduction in such networks of current, liquid, information, etc. It turned out that the properties are many – some real networks differ significantly from the properties of classical random graphs. The study of such parameters of complex networks as clustering, mediation or vulnerability are directly related to the theory of survivability , since it is these properties that determine the ability of networks to maintain their operability during a destructive effect on their individual nodes or edges (connections).

Despite the fact that the theory of complex networks includes various networks – electrical, transport, information, the greatest contribution to the development of this theory was made by the study of social networks. The term "social network" refers to the concentration of social objects, which can be considered as a network (or graph), the nodes of which are objects, and the links are social relations. The term was coined in 1954 by Manchester School sociologist J. Barnes in Classes and Gatherings in the Norwegian Island Parish. In the second half of the 20th century, the concept of "social network" became popular among Western researchers, and not only representatives of society, but also other objects that have social connections were considered as nodes of social networks. In the theory of social networks, such a direction as the analysis of social networks (*Social Network Analysis , SNA)*has been developed. Today, the term "social network" denotes a concept that has turned out to be wider than its social aspect. It includes, for example, many information networks, including the web space or social Internet networks.

Within the framework of the theory of complex networks, not only statistical, but also dynamic networks are considered, in order to understand the structure of which it is necessary to take into account the principles of their evolution [45].

In the theory of complex networks, there are three main areas: the study of statistical properties that characterize the behavior of networks; creating a network model; predicting the behavior of networks when changing structural properties. Applied research usually uses such typical network analysis characteristics as network size, network density, degree of centrality, etc.

One can speak of a "community structure" in a complex network when there is a network fragment – a group of nodes that have a high density of edges between themselves, while the density of edges between individual fragments is low. The traditional method for revealing the structure of communities is cluster analysis. Exist

dozens of acceptable methods for this, which are based on different measures of distances between nodes, weighted path indices between nodes, etc. In particular, for large social networks, the presence of a community structure turned out to be an integral property.

The loss of survivability of an information system can be caused by a break in links between its components, for example, when the most significant components are removed from the information space, i.e. those that have, say, the largest intermediation coefficient (*betweenness*). This coefficient for a particular network node is defined as the sum over all pairs of network nodes of the ratio of the number of shortest paths between them passing through a given node to the total number of shortest paths between them.

In the analysis of complex networks, as in graph theory, the parameters of individual nodes, the parameters of the network as a whole, and network substructures are studied.

3.1.1. Host settings

For individual nodes, the following parameters are distinguished:

— the input degree of connectivity of the node – the number of edges of the graph that are included in the node;

— the output degree of connectivity of the node – the number of edges of the graph that go out of the node;

is the distance from this node to each of the others;

is the average distance from a given node to others;

— eccentricity (*eccentricity)* — the largest of the geodesic distances (minimum distance between nodes) from a given node to others;

— mediation (*betweenness*), showing how many shortest paths pass through this node;

— centrality – the total number of connections of a given node in relation to others;

— a vulnerability considered as a level of network performance decline in case of deletion of a vertex and all edges adjacent to it.

The degree of connectivity k_i of a node i is the number of edges connected to that vertex.

Accordingly, the average degree of the entire network is calculated as the average of all k_i for all network nodes.

As noted above, in the case of directed networks, there are two types of node connectivity: output, corresponding to the number of edges outgoing from the given node, and input, equal to the number of edges entering the given node.

3.1.2. General network settings

To calculate network indices in general, the following parameters are used: number of nodes, number of edges, geodesic distance between nodes, average distance from one node to others, density – the ratio of the number of edges in the network to the possible maximum number of edges for a given number of nodes, the number of symmetrical, transitive and cyclic triads, network diameter is the largest geodetic distance in the network, vulnerability calculated as the maximum vulnerability of all network vertices, assortativity as a measure of correlation between degrees of nodes, etc.

There are several urgent problems of studying complex networks in terms of survivability, among which the following main ones can be distinguished:

— definition of network fragments (cliques, clusters), in which the nodes are more strongly interconnected than with members of other similar fragments;

— selection of network fragments (connectivity component) that are internally connected and not interconnected;

— finding jumpers, i.e. nodes, upon removal of which the network breaks up into unconnected parts.

3.1.3. Distribution of degrees of connectivity of nodes

An important characteristic of a network is the node degree distribution function $P(k)$, which is defined as the probability that a node i has a degree $k_i = k$, i.e. _ the degree distribution $P(k)$ reflects the proportion of vertices with degree k.

For oriented networks, there is a distribution of the outgoing half — degree $P^{out}(k^{out})$ and the half — degree of the input $P^{in}(k^{in})$, as well as a distribution of the total degree $P^{io}(k^{in}, k^{out})$. The latter specifies the probability of finding a node with input half degree k^{in} and exit noah half degree k^{out}.

Networks characterized by different $P(k)$, show very different behavior. $P(k)$ in some cases it can be a Poisson distribution ($P(k) = e^{-m}m^k / k!$, where m is the mathematical expectation, exponential ($P(k) = e^{-k/m}$) or power ($P(k) \sim 1/k^{\gamma},\ \ k \neq 0,\ \ \gamma > 0$).

An important feature of many real networks is the distribution of node degrees $P(k)$ according to a power law.

Networks with a power — law distribution of node connectivity are called *scale — free* . It is scale — free distributions that are often observed in real — life complex networks. With a power — law distribution, the existence of nodes with a very high degree is possible, which is practically not observed in networks with a Poisson distribution.

3.1.4. Path between nodes

If two nodes i and j can be connected using a sequence of m edges, then such a sequence is called a route (*walk)*between nodes i and j , and m is called the length of the route.

We say that nodes i and j are connected if there is a route between them. The connection relation is transitive, i.e. if node i is associated with node j , and j is associated with k , then it i is associated with k . In this case, a route whose beginning and end are in the same node, and all other vertices are used exactly once, is called a cycle.

The distance between nodes is defined as the length of the route from one node to another. Naturally, nodes can be connected directly or indirectly. A path between nodes d_{ij} is the shortest distance between them. For the entire network, you can introduce the concept of the average path as the average over all pairs of nodes of the shortest distance between them:

$$l = \frac{2}{n(n+1)} \sum_{i \geq j} d_{ij} ,$$

where n is the number of nodes ; d_{ij} is the shortest distance between nodes i And j .

Hungarian mathematicians P. Erdős and A. Rényi showed that the average distance between two vertices in a random graph grows as a logarithm of the number of vertices [46, 47].

In practice, the survivability of a communication network is defined as the probability of a path between any pair of nodes.

Some networks may be disconnected, i.e. there are nodes in them, the distance between which is infinite. Accordingly, the average path may also be equal to infinity. To take into account such cases, the concept of global network efficiency is introduced as the average inverse path between nodes, calculated by the formula

$$E = \frac{1}{n(n-1)} \sum_{i \neq j} \frac{1}{d_{ij}},$$

where the sum takes into account all pairs of nodes. This characteristic reflects the efficiency of the network in sending information between nodes (it is assumed that the efficiency in sending information between two nodes i And j inversely proportional to the distance between them).

The reciprocal of the global efficiency is the harmonic mean of geodetic distances:

$$h = \frac{1}{E}.$$

Since this formula eliminates the problem of discrepancy in determining the average path, this characteristic is better suited for graphs with several connected components.

The effective distance between two nodes is generally greater than the shortest distance.

Networks are also characterized by such a parameter as the diameter or the maximum length of the path, i.e. path equal to the maximum value of all d_{ij}.

3.1.5. Clustering coefficient

D. Watts and S. Strogatz in 1998 defined such a network parameter as the clustering coefficient [48], which corresponds to the level of connectivity of nodes in the network. This coefficient characterizes the tendency to form groups of interconnected knots, the so — called *cliques* . In addition, for a particular node, the clustering coefficient shows how many nearest neighbors of a given node are also nearest neighbors to each other.

The clustering coefficient for an individual network node is determined as follows. Let the nodes come out of k the edges that connect it with k other nodes, nearest neighbors. If we assume that all nearest neighbors are connected directly to each other, then the number of edges between them would be $\frac{1}{2}k(k-1)$, i.e. _ this is the number that corresponds to the maximum possible number of edges that could connect the nearest neighbors of the selected node. The ratio of the real number of edges that connect the nearest neighbors of a given node to the maximum possible (one at which all the nearest neighbors of a given node would be connected directly to each other) is called the clustering coefficient of the node i — $C(i)$. Naturally, this value does not exceed unity.

There is another way to calculate the clustering (transitivity) coefficient, based on the following formula:

$$C = \frac{3N_\Delta}{N_3},$$

where N_Δ is the number of 3 — cycles in the network, and N_3 is the number of connected 3 — components.

A 3 — cycle is defined here as a set of three nodes with edges between each pair of nodes. A connected 3 — component is a set consisting of three nodes, in which each node is reachable from another node, either directly or indirectly. Thus, in a 3 — component, the central node must be incident to the other two. The factor 3 was introduced taking into account the variants of different 3 — components for each 3 — cycle; this factor ensures the fulfillment of the inequality $0 \le C \le 1$. Then we get

$$N_\Delta = \sum_{k>i>j} a_{ij} a_{ik} a_{jk};$$

$$N_3 = \sum_{k>i>j} \left(a_{ij} a_{ik} + a_{ji} a_{jk} + a_{ki} a_{kj} \right),$$

where a_{ij} are the elements of the adjacency matrix A corresponding to the network, the sum is taken over all components of different nodes i, j and k only once.

The clustering coefficient can be determined both for each node and for the entire network. Accordingly, the level of clustering of the entire network is defined as the sum of the corresponding coefficients of individual nodes normalized by the number of nodes.

The difference between the two approaches to determining clustering is that, having averaged over the vertices, in the second case we get the same influence for each triangle in the network, and in the first case, an equal contribution for each node is taken into account.

This leads to different values of the clustering coefficient, because nodes with higher degrees are more likely to be part of more triangles than vertices with lower degrees.

The phenomenon of "small worlds" discussed below is directly related to the level of network clustering.

3.1.6. Mediation

The value of a node for the network is greater, the more paths it is involved in. Therefore , assuming that data is exchanged over the shortest paths between two nodes, one can quantify the value of a node in terms of betweenness , defined by the number of shortest paths passing through the node. This characteristic reflects the role of this node in establishing links in the network . The nodes with the most mediation play a major role in establishing links between other nodes in the network. Node m mediation b_m is determined by the formula

$$b_m = \sum_{i \neq j} \frac{B(i, m, j)}{B(i, j)},$$

where $B(i, j)$ is the total number of shortest paths between nodes i and j; $B(i, m, j)$ is the number of shortest paths between nodes i and j passing through node m.

Considering that shortest paths may not be known, and instead search algorithms are used to navigate the network, then the mediation (intermediate centrality) of a node can be expressed by the probability of finding it by the search algorithm.

The level of predominance of the largest intermediary in this case is determined in accordance with the formula:

$$CPD = \frac{1}{n-1} \sum_i (B_{\max} - B_i),$$

where $B_{\max}$ is the highest mediation level value in the network.

The predominance of the central node will be 0 for a clique and 1 for a star in which the central node is included in all paths.

3.1.7. Network elasticity and vulnerability

The opposite properties of network elasticity and vulnerability relate to the distribution of distances between nodes when individual nodes are removed. The elasticity of a network depends on its connectivity, i.e. the existence of paths between pairs of nodes. If a node is removed from the network, the typical length of these paths will increase. If this process continues long enough, the network will no longer be connected. R. Albert (Réka Albert) from the University of Pennsylvania (USA) in the study of attacks on Internet servers studied the effects that occur when a node is removed from the network, which is a WWW subset of 326,000 pages [49].

The average distance between two nodes, as a function of the number of nodes removed, hardly changed when nodes were randomly removed (high elasticity). At the same time, purposeful removal of nodes with the largest number of connections leads to the destruction of the network. Thus, the Internet is a highly elastic network with respect to the random failure of a node in the network, but highly sensitive to deliberate attack on nodes with high degrees of connections to other nodes.

One way to find critical network components is to look for the most vulnerable nodes [50] . If the performance of a network is related to its global efficiency, the vulnerability of a node can be defined as the performance degradation if the node and all its adjacent edges are removed from the network:

$$V_i = \frac{E - E_i}{E},$$

where E is the global efficiency of the original network, and E_i is the global efficiency after removing the node i and all edges adjacent to it.

The ordered distribution of nodes with respect to their vulnerabilities is related to the structure of the entire network . Thus , the most vulnerable node occupies the highest position in the network hierarchy. The measure of network vulnerability is the maximum vulnerability among all its nodes:

$$V = \max_i V_i .$$

3.1.8. Elitism Coefficient

Citation networks have been studied in scientometrics for a long time. It is known that influential researchers in certain areas form communities of a network type, expressed, for example, in the publication of joint works. This pattern is also observed in other real networks and reflects such a trend as good connectivity between hub nodes. This phenomenon, known as elitism (or rich — *club phenomenon*), can be characterized by the elitism coefficient introduced in [51]. An analysis of the topology of the web by S. Zhou and RJ Mondragon of the University of London showed that nodes with a high degree of outbound hyperlinks have more links to each other than to nodes with a low degree, while the latter have more connections with nodes with a large degree than among themselves. The study found that 27 % of all connections are between just 5% of the largest nodes, 60% are connections between the other 95% of nodes with the 5% of the largest, and only 13% are connections between nodes that are not in the top 5%.

The degree elitism k of the network G is a certain set of nodes with a degree greater than k, $\Re(k) = \{v \in N(G) \mid k_v > k\}$. The elitism coefficient of degree k is expressed as follows:

$$\varphi(k) = \frac{1}{|\Re(k)|\left(|\Re(k)|-1\right)} \sum_{i,j \in \Re(k)} a_{ij} ,$$

where the sum corresponds to twice the number of edges between the vertices in the "elite". This characteristic is similar to the clustering coefficient; it determines the proportion of links that exist between nodes with a degree greater than k .

3.1.9. Degree correlation of connected vertices

A significant number of structural and dynamic properties of the network is determined by estimating the correlation between the degrees of neighboring nodes. Such a correlation can be expressed in terms of the cumulative distribution $P(k, k')$, i.e. as the probability that an arbitrarily chosen edge connects a node of degree k with a degree node k'. The dependence between vertex degrees can be expressed in terms of the conditional probability that an arbitrarily chosen neighbor of a degree vertex k has a degree k' [52]:

$$P(k'|k) = \frac{\langle k \rangle P(k,k')}{kP(k)}.$$

At the same time $\sum_{k'} P(k'|k) = 1$. In the case of undirected networks $P(k,k') = P(k',k)$ and $k'P(k|k')P(k') = kP(k'|k)P(k)$.

If the network is oriented, then k is the degree of the previous node, k' is the degree of the next node, the values k and k' can be input, output or full degrees. In the general case $P(k,k') = P(k',k)$.

Values $P(k,k')$ And $P(k|k')$ formally describe the correlations of node degrees, but they are difficult to calculate experimentally, due to the size of the network and the low sampling power of nodes with high degrees. This problem can be solved by calculating the average degree of the nearest neighbors of nodes with a given degree k by the formula

$$S(k) = \sum_{k'} k'P(k'|k).$$

The indicator of correlation of degrees of connectivity makes it possible to single out separate classes of networks. If there is no correlation, then $S(k)$ it does not depend on the values of k, $S(k) = \langle k^2 \rangle / \langle k \rangle$. If $S(k)$ increases with increasing k, then nodes of large degrees tend to connections with nodes of large degrees, and the network is classified as assortative (hence the phenomenon of the "club of the rich"), while if is a $S(k)$ decreasing function of k, then vertices of large degrees tend to connections with vertices of small degrees, and the network is called disassortative [53].

In [53], the Pearson correlation coefficient was calculated for some real and simulated networks. It was found that, despite the models displaying the specific features of the structure (power — law distribution of the degree of connectivity of nodes, the “small world” property), most of them do not reproduce the assortativeness of a real network. For example, for the model of constructing a network with a power — law distribution of degrees of connectivity of Barabashi–Albert nodes [54], the value is $r = 0$. However, social networks tend to be assortative, while biological and technological networks are often disassortative.

It is known that assortative networks are less vulnerable to equiprobable attacks, and disassortative ones are less vulnerable to targeted attacks on hub nodes. Also, for example, synchronization network component states are faster in assortative networks. For example, when a contagious disease spreads, social networks should ideally be assortative: when controlling a small proportion of hub nodes, the network is divided into isolated connectivity components, which makes it possible to effectively control the spread of the infection.

3.2. Weak Ties Model

There is a class of complex networks that are characterized by so — called "weak" links. An analogue of weak social ties are, for example, relationships with distant acquaintances and colleagues. In some cases, these ties are more effective than "strong" ties. Thus, a group of researchers from the UK, the USA and Hungary obtained a conceptual conclusion in the field of mobile communications, which is that “weak” social ties between individuals are the most important for the existence of a social network [55].

For the study, calls were analyzed from 4.6 million mobile subscribers, which is about 20% of the population of one European country. This was the first time in the world practice when it was possible to obtain and analyze such a large sample of data related to interpersonal communication.

In a social network with 4.6 million nodes, 7 million social connections were identified, i.e. mutual calls from one subscriber to another and back, if return calls were made within 18 weeks. The frequency and duration of conversations were used to determine the strength of each social connection.

It was revealed that it is weak social ties (one or two callbacks over 18 weeks) that tie together a large social network. If these connections are ignored, the network will fall apart into separate fragments. If strong connections are not taken into account, then the connectivity of the network will be broken (Fig. 10).

It turned out that it is weak ties that are the phenomenon that binds the network into a single whole. It must be assumed that this conclusion is also valid for the web space, although research in this area has not yet been conducted.

1 2 3

Figure 10. Network structure:

1 – complete map of the social communications network; 2 – social network, from which weak ties were removed; 3 – a network from which strong ties are removed: the structure retains connectivity

3.3. Small world model

Despite the huge size of some complex networks, in many of them (in the web space in particular) there is a relatively short path between any two nodes – the geodesic distance. In 1967 psychologist S. Milgran, as a result of large — scale experiments, calculated that there is a chain of acquaintances, on average six long, between almost any two US citizens [56].

D. Watts and S. Strogatz discovered a phenomenon that is characteristic of many real networks, called the effect of "small worlds" (*Small Worlds)*[57].

Network structures corresponding to the properties of small worlds have the following typical properties: a small average path length relative to the network diameter (which is also typical for random networks) and a large clustering coefficient (which is inherent in networks with a regular structure).

When studying this phenomenon, they proposed a procedure for constructing a visual model of the network, which is inherent in this phenomenon.

To build a "small world" network, one should start with a regular cyclic lattice with N vertices, each of which is connected to k (in particular $k =$ 2) nearest neighbors in each direction. For each vertex, 2 k links are specified, where $N >> \log_2(N) >> 1$. Then each edge is reconnected to a random pair of vertices with probability p.

Under the condition, $p = 0$ an ordered lattice with a large number of cycles and large distances is obtained, and under the condition, $p \to 1$ the network becomes a random graph with short distances and a small number of cycles. In some average case, there are both short distances and a large number of cycles.

Three states of this network are shown in Fig. 11: a regular network, each node of which is connected to four neighbors, the same network, in which some "near" connections are randomly replaced by "far" ones (in this case, the phenomenon of "small worlds" occurs), and a random network in which the number of similar substitutions has exceeded a certain threshold.

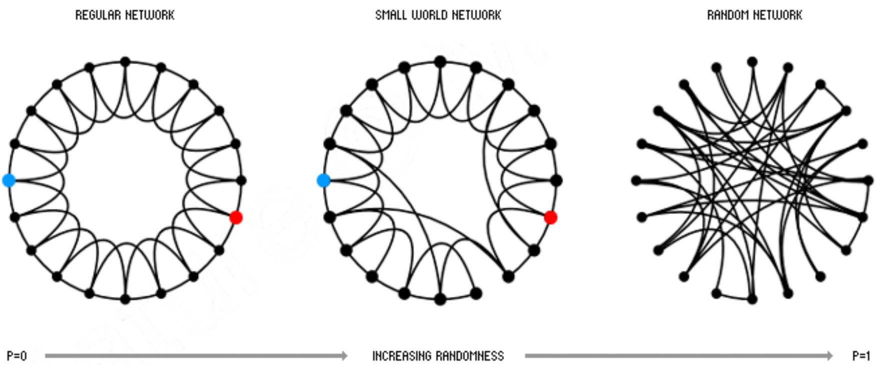


Figure 11. Watts — Strogatz Model

On fig. 12 shows graphs of changes in the average length of the path and the clustering coefficient of the artificial network by D. Watts and S. Strogatz on the probability of establishing "distant connections" (on a semi — logarithmic scale).

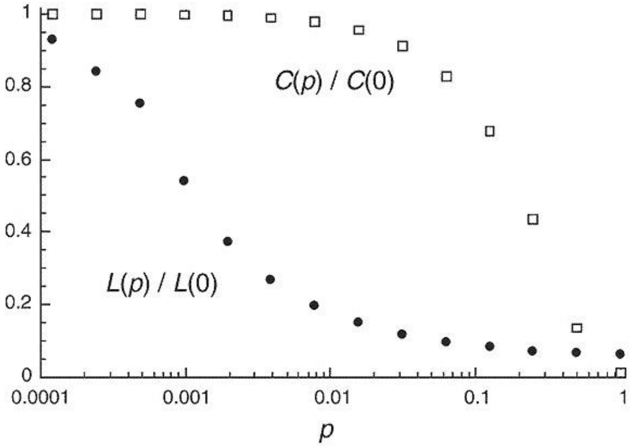


Figure Fig. 12. Dynamics of changes in the path length and clustering coefficient in the Watts–Strogatz model on a semi — logarithmic scale (the horizontal axis is the probability of replacing short — range links by far ones)

In reality, it turned out that it is precisely those networks whose nodes have simultaneously a certain number of local and random “distant” connections that simultaneously demonstrate the “small world” effect and a high level of clustering. Web space is also a network for which the phenomenon of small worlds is also confirmed.

These studies suggest that the dependence of the web space on large nodes is much more significant than previously thought; it is even more susceptible to malicious attacks. Related to the concept of "small worlds" is also a practical approach called "network mobilization", which is implemented over the structure of "small worlds". In particular, the speed of information dissemination due to the effect of "small worlds" in real networks increases by orders of magnitude compared to random networks, because most pairs of nodes in real networks are connected by short paths.

In addition, scalable, static, hierarchical "small worlds" and other networks are being studied quite successfully today; their fundamental properties such as resistance to deformation and percolation are investigated. It has recently been shown that a special class of networks, called "tangled" (eng. – *entangled networks*). They are characterized

maximum uniformity, minimum distance between any two nodes, and a very narrow range of basic statistical parameters. It is believed that entangled networks can be widely used in the field of information technology, in particular, in new generations of the web, allowing a significant reduction in network traffic.

3.4. System survivability branched structure

To assess the survivability of a system with a branched structure, it is necessary to establish a quality indicator, by changing which the degree of performance deterioration is determined in the presence of external excitatory influences. For branched structures, it is natural to choose the number of healthy branches as such an indicator. Since this number is a random variable, the distribution of the number of healthy branches must be found $\{P(t/S_i)\}$. On the vector of functions, $P(t/S_i)$ the functional is determined $\varphi\left(P(t/S_i)\right)$, which can be considered as an indicator of efficiency, by changing which the survivability can be assessed.

Consider the following performance indicators:
— *average number of healthy branches*

$$\varphi_1\left(P(t/S)\right)=\bar{N}(S)=\sum_{i=1}^{N_0}P_i(t/S)(N_0-i),$$

where $P_i(t/S)$ is the probability that the branches t are inoperable at the moment i; N_0 — the number of branches in a fully functional structure;

— *the probability of operability of the number of branches, not less than the specified*

$$\varphi_2\left(P(t/S)\right)=\sum_{i=1}^{i_0}P_i(t/S),$$

where i_0 is the maximum allowable number of unhealthy branches (the number of unhealthy branches is not less than $N_0 - i_0$).

Let us assume that it is subjected to repeated exposure k — tiered structure with branching coefficients on the tiers $r_1, r_2, ..., r_k$.

Within the framework of this model, it is assumed that only the nodes of the system fall within the scope of point NIs, and malfunctions of various nodes of the same tier are equiprobable. Since the total number of nodes in the system is $1 + r_1 + r_1 r_2 + ... \ + r_1 r_2 ... r_k$, we can write the following expression:

$$\gamma_0 + \gamma_1 r_1 + \gamma_2 r_1 r_2 + ... + \gamma_k r_1 r_2 ... r_k = 1,$$

where γ_k is the probability of a malfunction of the node k of the — th rank after a one — time NV. Various hypotheses about the vulnerability of elements are possible, which take into account the characteristics of nodes of different ranks, in particular, such a hypothesis is possible that

$$\gamma_i = \gamma = (1 + r_1 + r_1 r_2 + ... + r_1 r_2 ... r_k)^{-1}.$$

Then, based on the hypothesis of the "uniformity" of the tiers, we obtain

$$\gamma_i = \frac{1}{(k+1) r_1 r_2 ... r_i}, \quad i \geq 1; \quad \gamma_0 = \frac{1}{k+1}.$$

In a more general case, one can

$$\gamma_i = a_i \gamma_{i+1}, 0 \leq i \leq k-1, a_i \geq 1.$$

Then we get

$$\gamma_k = \gamma, \quad \gamma_i = a_i a_{i+1} ... a_{k-1} \gamma, \quad 0 \leq i \leq k-1, \ \gamma = \left(\sum_{n=0}^{k} \prod_{i=n}^{k-1} a_i \prod_{j=1}^{n} r_j \right)^{-1}.$$

3.5. Modeling the destructive network impact

Like the network of terrorists, the restoration of which after a destructive impact is described in [3], the information system is also a dynamic system, the restoration of which after the destruction of the best "intermediaries" is carried out through latent links with other components of the information space. After the information system is divided into isolated fragments, it can "use" these links and quickly restore connectivity (Fig. 13).

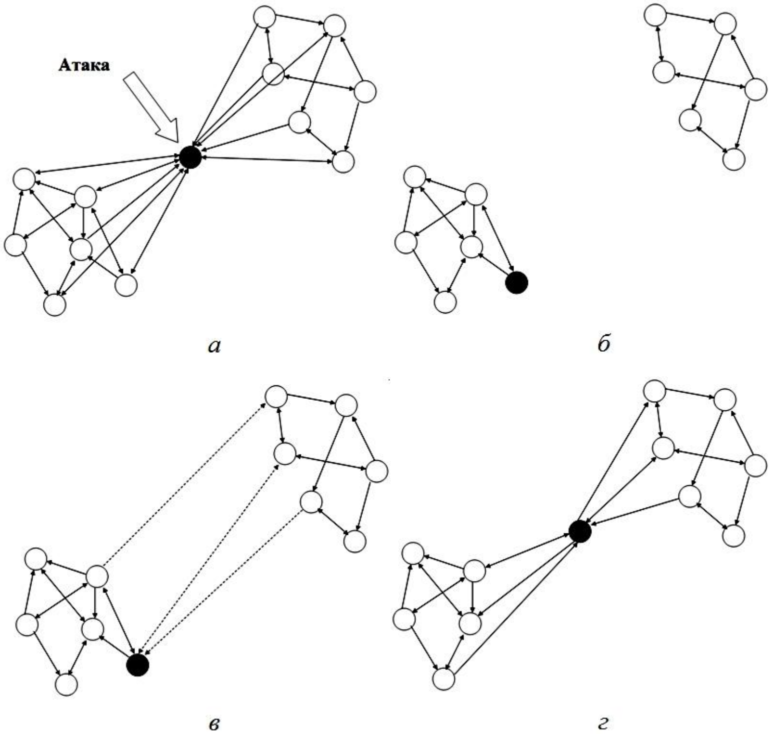


Figure 13. Restoring the network structure by choosing a new one "intermediary": a – attack on the network; b – disconnected network with removed intermediary; c – the revival of hidden (latent) connections; d – network connectivity restored

The reunification of parts of the network will not take place if none of the pairs of components that survived the destructive impact can find hidden connections between themselves (perhaps not directly, but through other components of the information space). In this case, the impact of disconnection on the performance of the information system depends on whether the again disconnected parts can get interconnections, the lack of which is observed in this part of the information system.

If a part of the information system was close to self — sufficiency, then it continues to function independently. Otherwise, it stops functioning until a new connection is formed. If one of the connections is successful, then its initiator becomes a new "intermediary" that connects the two parts of the network.

4. CRITICAL LEVEL SYSTEMS SURVIVABILITY

If there are too few edges in the network, i.e. the average degree of its vertices is small enough, then it contains many isolated nodes and clusters of a few nodes. When adding edges to such a network (for example, establishing new links, quoting, etc.), small clusters are combined into large ones.

Upon reaching a certain critical level characterizing the network flow threshold, most of the vertices will be combined into a cluster, the size of which is comparable to the number of network nodes, i.e. e. the entire network will become conductive.

This effect, a phase transition, is described in the literature as “the appearance of a giant cluster” [58].

When studying the survivability of the network structure, the main interest is the opposite effect – the transition from a giant cluster to a sparse network as a result of destructive influences, expressed in the removal of network elements – edges (or nodes). In this case, a functional failure is considered as the removal of an individual element. In this case, a fairly accurate analogy arises with the percolation limit (or percolation threshold) known in percolation theory, which is associated with a phase transition.

The existing mathematical models of such behavior of network structures (diode networks, networks of terrorists, etc.) under destructive action are reduced to a percolation problem [59–61]. As a rule, within the framework of this problem, a lattice of N conducting elements is considered, after which the conductivity of the entire lattice is calculated with the removal of its individual elements. The conductivity threshold is considered as an analogue of the survivability threshold.

On fig. 14 shows an example of the typical behavior of such models. With an increase in the number of functional failures k there is a decline in the number of functions performed. In the general case, at the point, which we will call the "critical level of survivability", there is a sharp drop (phase transition of the second kind).

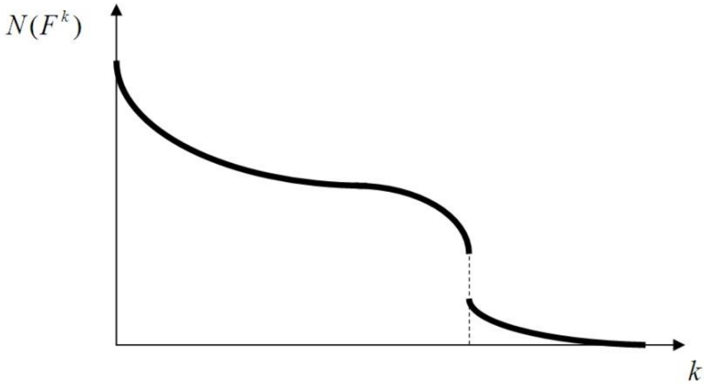


Figure 14. Dependence of the number of functions performed (vertical axis) on the number of functional failures k (number of elements that failed)

4.1. Phase transitions

The concept of phases is found in many phenomena, here are a few examples. Three phases are known for water: liquid, solid (ice) and gaseous (steam). Each of them is characterized by its own parameter values. It is essential that when external conditions change, one phase (ice) passes into another (liquid). Another object is a ferromagnet (for example, iron or nickel). At low temperatures $t < T_c = 360\ ^\circ\mathrm{C}$, nickel is a ferromagnet; when the external magnetic field is removed, it remains magnetized. At temperatures above T_c this property is lost, when the external magnetic field is turned off, it goes into a paramagnetic state. Here, as in the previous example, the existence of different phases, paramagnetic and ferromagnetic, is clearly seen. When the temperature changes, a phase transition occurs.

In [60], an example from percolation theory is given. Randomly cutting out edges from the grid, in the end, when the concentration of the remaining edges p becomes less than a certain value p_c, it will no longer be possible to pass through the lattice “from end to end”. Thus, the grid from the "leakage" phase will go into the state of the "non — leakage" phase.

From these examples, it is clear that for each of the systems considered, there is a so — called order parameter that determines which of the phases the system is in. In ferromagnetism, the order parameter is the magnetization in a zero external field; in percolation theory, it is the network connectivity.

Phase transitions are of several types. A phase transition of the first kind is characterized by the fact that several phases can simultaneously exist in the system. For example, at 0°C, ice floats in water. If the system is in thermodynamic equilibrium (there is no drive and heat removal), then the ice does not melt and does not grow. For phase transitions of the second kind, the existence of several states simultaneously is impossible. A mesh with randomly cut edges is either connected or not.

Decisive in the creation of the theory of phase transitions of the second kind, the beginning of which was laid by L.D. Landau, was the introduction of the order parameter η as a distinctive feature of the phase of the system. In one of the phases, for example paramagnetic, $\eta = 0$, and in the other, ferromagnetic, $\eta \neq 0$. For magnetic phenomena, the order parameter η is the magnetization of the system.

To describe phase transitions, a certain function of the parameters that determine the state of the system, , is introduced $G(\eta, T, ...)$. In physical systems, this is the Gibbs energy. In each phenomenon (percolation, a network of "small worlds", etc.), this function is defined in its own way. The main property of this function, the first assumption of L.D. Landau – in the state of equilibrium, this function takes the minimum value:

$$\frac{\partial G}{\partial \eta} = 0, \quad \frac{\partial^2 G}{\partial \eta^2} > 0.$$

In physical systems one speaks of thermodynamic equilibrium, in the theory of complex networks one can speak of stability. Note that the minimality condition is determined by varying the order parameter.

4.2. The problem of percolation theory

One of the important characteristics of complex networks is the possibility of current, liquid, information (traffic), etc. flowing along their edges. For the first time, the problem of percolation (English *percolation* –

seepage, flow) was formulated in 1957 in the work of S.R. Broadbent (SR Broadbent) and J.M. Hammersley (JM Hammersley) [61]. Subsequently, a whole area of research was developed, called the theory of percolation, which has numerous applications in practice. It turns out that many questions that arise in the analysis of the survivability of information segments are also directly related to the theory of percolation.

The theory of percolation faces many questions that go beyond the standard framework of discrete mathematics and probability theory [59].

The simplest formulation of the percolation theory problem is as follows. A lattice of bonds is given, the random part of which p is "black", conducting, and the rest is "white", not conducting the flow. It is necessary to find such a minimum concentration p_c of "black" bonds, at which there is still a connected path along the "black" bonds through the entire lattice , i.e. _ such a concentration that the lattice as a whole conducts the flow.

At $p = 0$ all lattice bonds are “white” — the lattice does not conduct the flow. With an increase in the concentration of "black" – conducting bonds, at , $p = p_c$ a percolation, penetrating cluster of "black" bonds appears in the lattice, connecting the opposite edges of the grid. When the grid size tends to infinity, the size of this cluster is also infinite, in connection with which the term “infinite cluster” was introduced. Other sets of interconnected bonds of finite size are called finite clusters.

When passing through the percolation threshold, i.e. when an infinite cluster appears, the properties of the system that characterize it as a whole change dramatically. If, for example, the "black" bond conducts current, but the "white" bond does not, then the conductivity of the entire system $R \sim 1/G$ nearby p_c decreases sharply. Considering the logarithmic scale in Fig. 15, it is clear that near the percolation threshold, the throughput of the network, with a very small decrease, p can drop sharply.

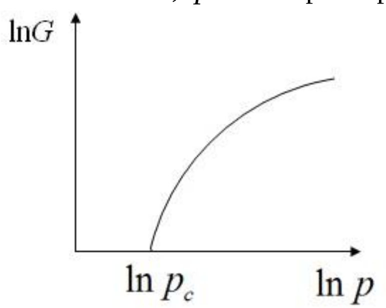


Figure 15. Conductivity of the system near the percolation threshold (on a logarithmic scale)

It would seem a simple task of determining p_c, but "did not succumb" to the exact methods of probability theory. With rare exceptions, p_c it is not possible to calculate analytically – numerical simulation is necessary.

At present, many important generalizations of the percolation problem are known, for example, cases are considered when "non — conducting" bonds conduct, but are much worse than conductive ones; we can talk about the values of conductivity for different bonds; one can consider unidirectional "diode bonds", etc.

The tasks solved in the framework of the theory of percolation and analysis of complex networks include such as determining the limiting level of conductivity (capacity), changing the length of the path and its trajectory (tortuosity, parallelism) when approaching the limiting level of conductivity, the number of nodes that need to be removed to break network connectivity.

4.3. Characteristics of percolation networks

There are important characteristics of the edges of networks, along which, for example, the passage of current, liquid, information (in the cases of interest to us) occurs.

Here are some of them:

$P_\infty(p)$ — the probability that a randomly selected node belongs to an infinite cluster connecting two opposite —

Opposite sides of the network – this characteristic makes sense above the percolation threshold when there is an infinite cluster;

$S(p) = \dfrac{\sum_s s^2 n_s(p)}{\sum_s s n_s(p)}$ – average cluster size (English – *mean cluster size*), where n_s — the number of clusters of s nodes per lattice node –

this characteristic makes sense below the percolation threshold, when all clusters existing in the system have a finite size;

$\xi(p)$ is the correlation length characterizing the decay rate of correlations in the lattice. To introduce it, it is necessary to define the pair correlation function $G(r, p \approx p_c) \equiv G(|r_i - r_j|, p) = = \left\langle g(r_i, r_j) \right\rangle$, where is the function $g(r_i, r_j)$ on the radii of the node vectors i and j is equal to one if the nodes are connected by "black" links belonging to one finite cluster, and equal to zero in all other cases, and the angle brackets mean averaging over all nodes. At , $r \to \infty$ the correlation function $G(r, p)$ decreases exponentially, and the characteristic scale in this case is precisely the correlation length $G(r, p):\ \exp[-r/\xi(p)]$.

Note that at , $p \to p_c$ the correlation length $\xi(p)$ diverges, which is in full agreement with the qualitative concepts. The closer to the percolation threshold $(p \geq p_c)$, the fewer "black" conductive bonds, the more significant each part of the cluster connecting two "infinities", the breakage of one of them can affect far from it. At the very threshold of percolation, one dangling bond can destroy the entire conducting path, called an infinite cluster. The structure of an infinite cluster is not simple; at the very threshold of flow, it is a fractal object consisting of links included in this cluster, with dimension $d_f(d=2) = d - \beta/\nu \approx 1{,}896$, $d_f(d=3) \approx 2{,}54$.

Other characteristics of an infinite cluster are also introduced – the backbone of the cluster (eng. *backbone)* — the part of the cluster along which the transfer occurs, as well as cut off "dead ends" (eng. *dead ends*), and many others.

In a more developed and realistic theory of percolation, “white” bonds are also considered conductive, with a conductivity much lower than “black” ones (it is customary to say that it is one times $1/h$ less, where $0 < h \ll 1$). With the introduction of a parameter, h the percolation theory can be formulated in terms of the theory of second — order phase transitions, one of the most complex branches of theoretical physics.

The description of the conductivity of a percolation network can also be formulated in terms of the theory of second — order phase transitions. For definiteness, we will talk about the so — called effective conductivity σ_e, which characterizes the conductivity of the system as a whole. The role of the order parameter in this case is played by σ_e, and the proximity to the critical point $\tau = (p - p_c)/p_c$, where p_c is the percolation threshold. The corresponding scaling relation for effective conductivity σ_e has the form

$$\sigma_e = h^{\frac{t}{\varphi}} f(\tau / h^{\frac{1}{\varphi}}),$$
$$f(z) \sim \begin{cases} z^t, & z \to +\infty, \\ \text{const}, & z \to 0, \\ |z|^{-q}, & z \to -\infty. \end{cases}$$

In these formulas t, and q are the so — called critical conductivity indices, $\varphi = t + q$.

It should be clarified that the universal scaling behavior σ_e takes place only near the percolation threshold, i.e., at $|\tau| \ll 1$.

4.4. Percolation on random networks

Percolation can also be considered on random networks of Erdos — Renyi, Watts — Strogats, etc. In this case, instead of an infinite percolation cluster, one speaks of a giant connected *component* .

For a random Erdos–Rényi network of N nodes, it is known that the percolation threshold $p_c \approx 1/N$, i.e. that leakage occurs when the average node degree is $\langle k \rangle \geq 1$.

M. Newman and D. Watts [58] considered the problem of percolation on "small world" networks. In this case, a modification of such networks was used. In contrast to how the "small world" was introduced in [57], new connections, called *shortcuts* , were additionally thrown into the initial network, all old connections between neighbors were not interrupted, remaining in their places. As in the standard percolation theory, a characteristic length can be introduced (in percolation theory it is called the correlation length):

$$\xi = \frac{1}{(\varphi kd)^{1/d}},$$

where φ is the probability of encountering a redirected link (*shortcuts*), φN is the number of these links, N is the number of network nodes, k is the number of neighbors, d is the dimensionality of the network. The meaning of this value is the distance between the ends of different thrown connections.

For the shortest path l (the middle one, of course) there is a scaling function:

$$l = \frac{N}{k} f\left(\frac{N}{\xi}\right),$$

$$f \;\; z \ll 1 \; \to \text{const}, \qquad f \;\; z \gg 1 \; \to \frac{\log z}{z}.$$

Since $\xi \sim 1/(\varphi k)^{1/d}$, the asymptotics of the scaling function means for $z = N/\xi \ll 1$, that $N\varphi^{1/d} \ll 1$, i.e. that the number of shortcuts *is* very small. Then

$$l = \frac{N}{k} f\left(\frac{N}{\xi}\right) \sim N,$$

those. the network is a "big world", the average distance between two nodes increases in proportion to the number of nodes.

If $z = N/\xi \gg 1$, i.e. $N\varphi^{1/d} \gg 1$, then

$$l = \frac{N}{k} f\left(\frac{N}{\xi}\right) \sim \log\left[N(\varphi kd)^{1/d}\right] \sim \log N,$$

the network is a "small world", the distance between nodes increases much more slowly, as the logarithm of the number of nodes.

Let us now return to the question of the percolation properties of such networks. To study the problems of system survivability, several questions are of particular interest, for example, what will happen if a certain proportion of $q = 1 - p$ nodes falls out (does not conduct information, current, etc.), which is equal to the critical value $q_c = 1 - p_c$ when there is still a giant connected component in the network , i.e. when a significant number of network nodes are still interconnected? It is known that the percolation threshold p_c (Np_c is the number of unbroken bonds) is related to the share of transferred bonds φ as follows:

$$\varphi = \frac{(1 - p_c)^k}{2kp_c\left[1 + kp_c(1 - p_c)^k\right]}.$$

On fig. 16 shows the dependence of the percolation threshold p_c on φ, at which there is still a giant connected component. As can be seen, during the flow, the smaller the transferred share of links, the greater the number of nodes must be transferred.

Percolation on scale — free networks with distribution of node degrees $P(k) \sim k^{-\gamma}$ has its own specifics, different from percolation in “small worlds”. The expression for the percolation threshold p_c is different for different ranges of the parameter γ. For example, when $\gamma > 3$:

$$q_c = 1 - p_c = 1 - \frac{1}{\frac{\gamma - 2}{\gamma - 3} k_0 - 1},$$

where the degree of the node lies in the range $k_0 \le k \le K_0$.

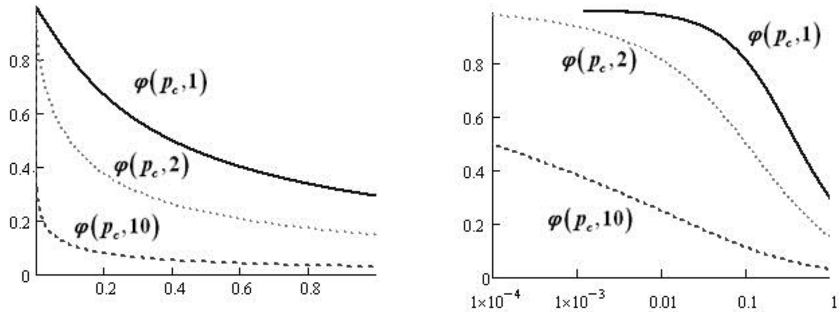


Figure Fig. 16. Dependence of the percolation threshold p $_c$ on the share of transferred bonds φ: from top to bottom – the number of neighbors is one, two and ten. On the right side, the x — axis is in a logarithmic scale

When $k_0 = 0$ the expression for q_c is simplified $q_c = 4 - \gamma$. Thus, if the proportion of remote nodes is greater than $4 - \gamma$ — there is no giant cluster, i.e. we can assume that the network is destroyed. Details about this and what happens in other ranges of the parameter γ can be found, for example, in [62].

4.5. Percolation theory and modeling attacks on networks

The presented results are quite suitable for modeling attacks on such a global network. It is known that Internet nodes are connected by communication lines with different bandwidth. Separate segments of the Internet are connected to traffic exchange points, there are backbone *channels*. Simulation of attacks, i.e. ruptures of individual connections or the removal of individual, as calculations show, very few nodes can disrupt such connectivity and give terrorists great chances. Finding "weak" sections of the global network, designing redundant nodes and channels require accurate calculations based on the theory of percolation.

It should be noted that above all the time it was about the disabling of elements (nodes or links) with a q random probability. At the same time, it is possible, for example, to cut out nodes both in the standard percolation problem and in other types of complex networks, in a purposeful way, choosing such nodes, when cutting out of which the network collapses as quickly as possible. On the Internet, such a targeted cutting of nodes (servers) is called a "planned attack". At the same time, disabling about 1% of targeted nodes reduces the performance of the entire network by half [63].

It is also possible to generalize the percolation problem to the so — called diode or directed *percolation*. In this modification, some of the well — conducting ("black") bonds in the two — phase version of percolation are replaced by "diode" ones, which pass only in one direction.

Individual websites (nodes) and hyperlinks in the web space form a diode percolation network, since according to the HTTP protocol it is not provided that hyperlinks leading from web pages lead in the opposite direction. A relatively small number of websites (for example, web resource directories) contain a large number of hyperlinks, while most websites contain very few links to other web resources. If we assume that an attack will be made on websites that have the largest number of links to other websites, then, depending on the power of such an attack, the connectivity of this diode network may be broken. As a result, search servers will not be able to cover some part of the web space, which will automatically transfer the latter to the "hidden" web zone [64].

5. INFORMATION FLOWS AND SURVIVABILITY ISSUES

5.1. Information flows

Information systems as network structures in the information space consist of separate elements that form information flows in the dynamics of their evolution (appearance, development, modification, destruction). Therefore, the survivability of information systems directly depends on the properties of information flows.

To study modern information flows on the Internet, i.e. message flows that are published on the pages of websites, social networks, blogs, and the like, a fundamentally new toolkit should be used, because the classical methods of generalizing information arrays (classification, phase enlargement, cluster analysis, etc.) are not always able to adequately reflect the state of the dynamic component of the information space. In this case, we are talking not so much about the analysis of documentary arrays of fixed sizes, even very large ones, but about generalizing the dynamic flow of hypertext data.

Of course, most of the information that is presented on the Internet finds its consumer. However, if we consider the entire set of online publications as a certain generality in relation to a specific user (or group of users), then we can see a number of problems associated with the completeness, relevance and efficiency of obtaining data. Searching, filtering, collecting information on the Internet requires sufficient staff qualifications and, unfortunately, all the features of the information structure of the network and the presentation of data in it cannot be taken into account. This, in turn, leads to the fact that single samples of information from the web space cannot be considered representative.

At the same time, the information flow that is “consumed” by a specific user, as a rule, has a pronounced subject orientation, which is characterized by the area of his interests. The search and processing of information in manual mode is a rather laborious, and most importantly, lengthy process, which most often does not the desired result. The solution of the problem in practice is possible by creating automated systems for collecting, filtering and analyzing information, the so-called intelligent intermediaries between the user or corporate information system and the Internet. Such a system should collect and select information from Internet and create a document database specified by the user's subject area, i.e. perform the functions of integrating information flows. Loading information into the database should be accompanied by its classification and structuring. For subsequent information and analytical work, the user should be provided with effective means of navigation, search and generalization of information that is stored in the corresponding dynamic documentary database.

The current level of development of the information space determines the interest in approaches based on the understanding of information as a measure of the orderliness of a certain system and, accordingly, in statistical methods for its processing. To organize effective communication in networks today, one has to constantly return to the mathematical origins of information theory, the concepts of entropy, Shannon's theory, Boltzmann's equations, etc., broad prospects for the use of a powerful apparatus of mathematics and physics in solving information-theoretic problems [65].

For a formal description of information flows, we introduce some general assumptions for the entire subsequent presentation. Let us give a definition of the information flow [66], which corresponds to the classical definition from information theory.

Consider a segment (a, τ) of the real axis (time axis), where $\tau > a$. Let's assume that in this period of time, in accordance with certain patterns, a certain number of information documents are published on the network - k. On the time axis, the moments of publication of individual documents will be denoted as $\tau_1,\ \tau_2, ...,\ \tau_k\ \ (a \le \tau_1 \le \tau_2 \le ... \le \tau_k \le \tau)$. An information flow is a process $N_\alpha(\tau)$ whose implementation is characterized by the number of points (documents) that appeared in the interval (a, τ) as a function of the right end of the segment τ. According to this definition, the implementation of the information flow is a non-decreasing step function that is always integer $N_\alpha(\tau)$.

The above definition on local time domains corresponds to reality, but does not take into account such an effect as information aging, which contradicts the "accumulative" ability of the information flow $N_\alpha(\tau)$ over long time intervals.

So, a certain information flow takes into account only the number of information messages, regardless of their content. In the general case, determining the content and subject matter of individual documents is a rather subjective process. For rigorous modeling of thematic information flows, models are used that distinguish documents by individual words or phrases (usually they are called terms, from English - *Terms)*.

The tasks of monitoring large volume information flows in computer networks, their adaptive aggregation and generalization are complicated by the lack of standard methods and solutions, the incompleteness of existing technological approaches. Currently, research on the problems of analyzing large volume information flows in computer networks is most often of a highly specialized nature. At the same time, the experience of creating and implementing corporate information systems indicates the need to create and implement documentary information repositories to support scientific research, obtain various analytical information, and navigate large volumes of documentary information flows.

When modeling these processes, the methods of nonlinear dynamics, the theory of cellular automata and self-organized criticality are used. When modeling information flows, the structural relationships between the arrays of documents included in them are studied. Today, in this case, fractal analysis is increasingly being used, an approach based on the properties of preserving the internal structure of arrays of documents with changes in their size or scale of consideration. Information theory, which previously found its main application in the field of data transmission, is also becoming useful for the analysis of text arrays dynamically generated in networks.

It is envisaged that news messages have the property of aging, i.e. lose their relevance over time. The entire information space can, with a sufficient degree of convention, be divided into two components - stable and dynamic, which have very different characteristics of their development. In particular, the information aging process in the well-known Barton - Kebler model is described by an equation consisting of two components:

$$m(t) = 1 - ae^{-T} - b\,\mathrm{e}^{-2T},$$

where $m(t)$ is a part of useful information in the general flow after time T; the first subtrahend corresponds to stable resources, and the second - to dynamic, news. This equation also fully corresponds to the volumes of information that are formed in the information space on certain topics, which appear and disappear from time to time. The stable component of the information space contains information of a "long-term" plan, while the dynamic component includes resources that are constantly updated. Some part of the latter component subsequently “merges” into the stable component, however, most of it “disappears” from the information space or falls into its segment, the so-called hidden part, which is not available to users using conventional information retrieval systems (IPS).

In a traditional network information retrieval system, the information space, consisting of stable and dynamic parts, which is indexed using this IPS, changes its content after a while: some documents are placed in the stable part in the form of archives, while others disappear. In this case, when accessing the IPS, the user finds documents relevant to the request from the stable part, links from the dynamic part that are outdated, and does not find anything from the updated dynamic part.

Navigation in information from the dynamic part of the information space requires the use of an intermediary system between the user and the information space. Such an intermediary (or news agent) can perform the work of collecting and selecting information. The principle of indexing to be carried out by this intermediary is slightly different from indexing by traditional search engines: not all the content of the information space, but only its dynamic part.

Information does not appear in the information space by itself. It is published, posted on websites, pages of social networks, nodes of peer-to-peer networks, and the like. In the future, such information resources will be called information sources.

It should be noted that the content monitoring system InfoStream [67], which is one of the practical “polygons” of this study, consumes a polythematic information flow from more than 5000 sources. On Fig. 17 shows a fragment of the dynamics of this stream (the vertical axis is the number of documents) in terms of time (days) and the most productive sources. This graph clearly shows weekly fluctuations in the volume of information flows.

At present, the powerful capabilities of the Internet give rise to the problem of optimizing the composition and number of sources that can be used by a corporate information system to provide the required quantity and quality of documents, satisfying meeting the needs of users. In this regard, the issues of ranking and selection of sources of news information are relevant - websites that need to be accessed through one interface both in the search mode and in the modes of revision and analytical generalization.

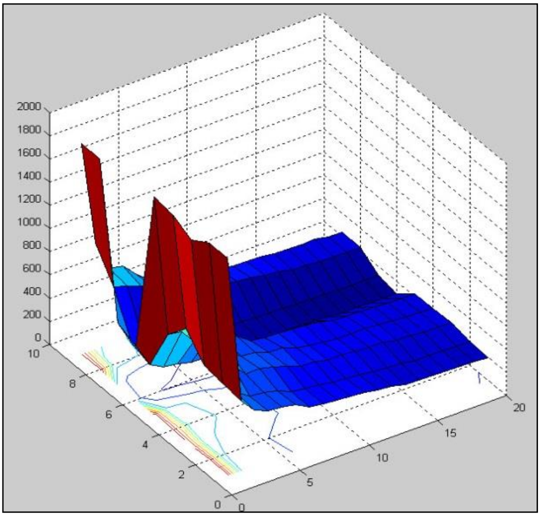


Figure 17. Dynamics of publication of information by the most rated sources (obtained using the InfoStream system)

A large number of scientific papers and practical developments are devoted to the principles of ranking web documents. Link ranking of websites today is a separate area of Internet business - SEO (*search engine optimization*). At the same time, much less attention is paid to the issues of ranking and selection of information resources, taking into account their content, volume and stability of publications. Of course, the main criterion for choosing sources for such news monitoring systems is their content. The distribution of sources by content that corresponds to the thematic needs of corporate users satisfies Bradford's law , respectively, when selecting sources, their ranking according to the degree of relevance to the topic must be taken into account.

However, the implementation of such a choice leads to some difficulties. In practice, such ranking is carried out by experts by estimating the number of documents relevant to a predetermined package of thematic queries that are addressed to a fragment of a database composed of documents from the analyzed source. And this inevitably leads to a certain subjectivism with all the ensuing consequences.

Therefore, it seems promising to supplement the traditional approach with more objective and more rigorous methods that allow optimizing the process of forming the information base of information flow integration systems.

On Fig. 18 shows the distribution graph (on a semi-logarithmic scale) of the number of documents published by sources scanned by the InfoStream system, which are ranked by the parameter - the number of documents published by the source.

The distributions below refer to the March 2008 document set of over 1.2 million documents from over 2,500 public web sources. The central part of the graph is well approximated by a straight line, which indicates that the presented dependence is close to hyperbolic (i.e.

on the operation of the generalized Zipf law). On fig. 19 shows the total number of documents that are covered by the monitoring system depending on the sources taken into account in it, also ranked by the number of published documents. Since Zipf's law allows the distribution density to be approximated by a hyperbolic dependence of the form a/x, then the distribution function for the number of documents

$$f(x) \sim \int \frac{a}{x} dx = a \ln x + C,$$

with a good approximation is described by a logarithmic law.

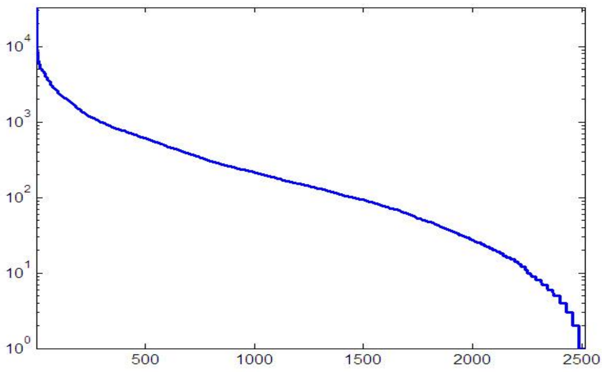


Figure 18. Distribution of the number of publications (y-axis) by ranked list of sources (abscissa)

The given dependency made it possible to construct a criterion for selecting the necessary part of the sources for different corporate applications from the general list that meet the needs of users.

As already noted, the citation of individual documents and websites today is one of the main criteria for assessing the ranks of documents in online search engines (PageRank, HITS, TrustRank, TIC, etc.). The idea of evaluating the citation level made it possible to implement one of the first models of the dynamic part of the web space [68]. Note that the assessment of the level of the source of information as "author" mainly by the number of websites from which hyperlinks are made to it is fully consistent with the Salsa algorithm proposed by Moran and Lempel.

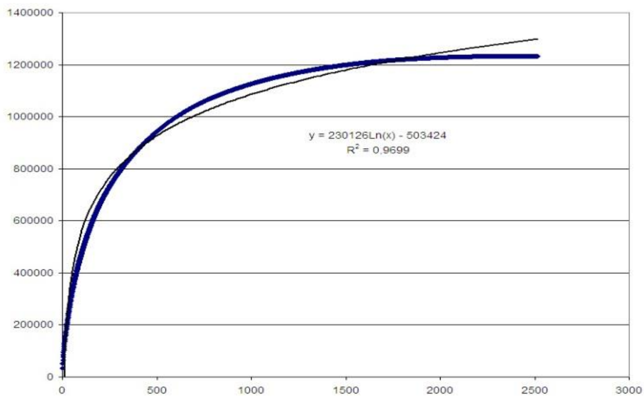


Figure 19. Dependence of the number of publications in the monitoring system (y-axis) from sources ranked by the number of documents (y-axis)

If we assume that all sources make the same contribution in terms of the number of published documents, then this relationship could be linear and expressed by the formula:

$$f_{lin}(n) = n\frac{f_{\max}}{N},$$

where $f_{\max}$ is the maximum volume of covered documents; N is the total number of sources; n — number of the current source.

Obviously, the deviation of the real dependence from the linear one first increases, and then decreases to zero. We will call the number of sources threshold n_p if the value of the real dependence deviates as much as possible from the reduced linear

$$n_p = \arg\max_n \ \{ f(n) - f_{lin}(n) \}.$$

On Fig. 20 shows the dependence n_p on various values of N, i.e. when N the most productive sources are selected.

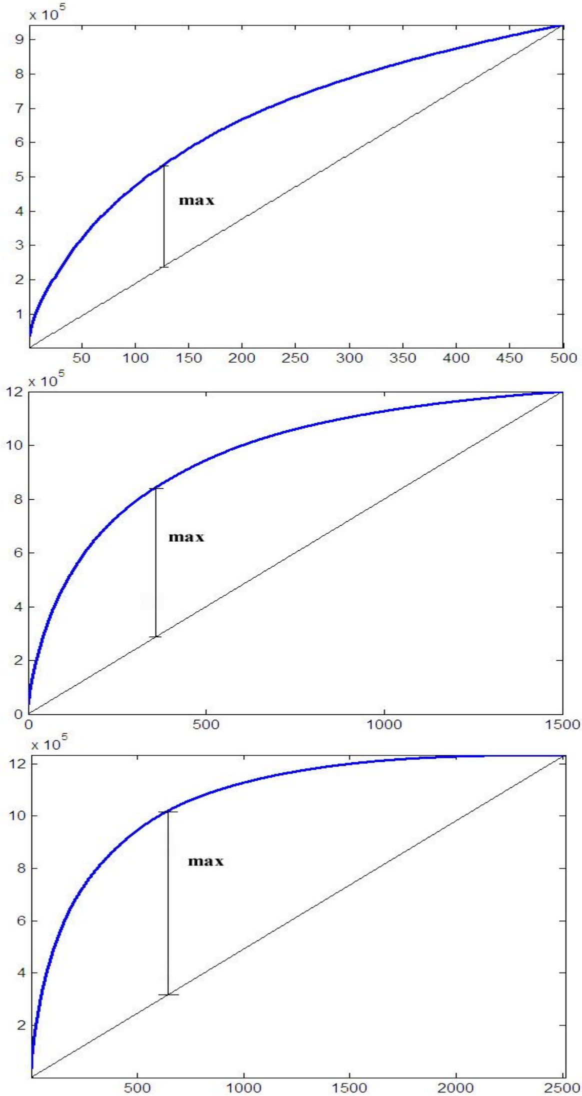


Figure 20. Number of publications in the monitoring system (y-axis) when new most intense sources are connected (abscissa)

It is interesting (and fully consistent with the nature of the function $f(n)$ that the values n_p depend almost linearly on N (Fig. 21): $n_p \sim 0,24N$, at the same time, the number of covered documents corresponding n_p to the maximum number of sources (2514), (Fig. 22) reaches 80 percent of $f_{\max}$.

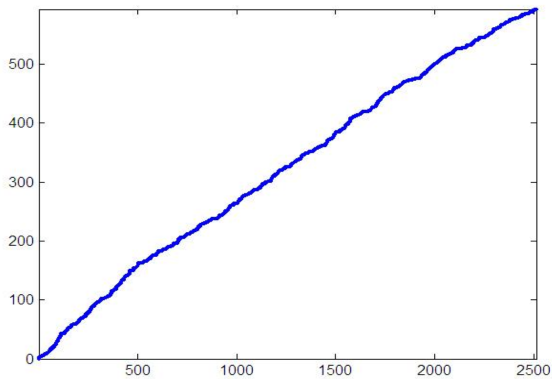


Figure 21. Threshold dependency (y-axis) from the initial number of sources (abscissa axis)

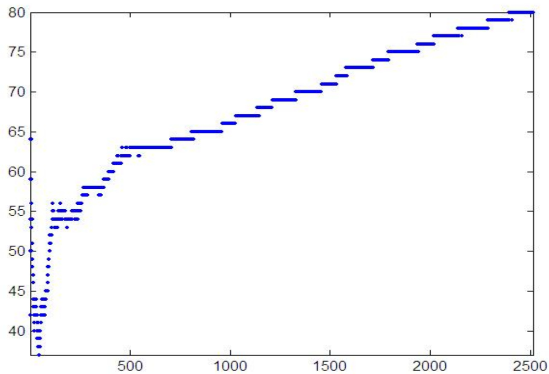


Figure 22. Dependence of the specific number of documents, covered by the system (y-axis), from a change in the initial number of sources (abscissa axis)

Note that the constructed dependence satisfies the Pareto principle: approximately 20 % of the most productive sources publish 80% of documents.

A special place in the study was occupied by the study of meaningful duplication of information. At the same time, it should be noted that the percentage of documents that are duplicated in meaning in the InfoStream monitoring system is much less than in the entire news web space. This is due to the selection of sources for scanning, which do not include many news integrators.

As already noted, one of the main features of news information is the presence of a large number of messages that duplicate each other. In addition, all mass media (media) will write about an event of world significance, and, most likely, on one of the first pages. The consumer (with the exception of some specific areas of analytical research of the information space), however, wants to receive one message for each event importance in modern technologies. In particular, the task of selecting the most original sources, which allow (at least statistically) to exclude not only formal, but also substantive duplication of information, becomes relevant. Duplication of messages on websites depends on various reasons, so the measurements for the list of sources ranked by the number of publications show a different level, and the information is not visual. At the same time, the study of the authors shows a steady trend: the more productive the source of information, the more it contains borrowings from other sources.

5.2. Information flow modeling

The survivability of information systems as thematic information flows is determined by their composition, structure and content (content). Analysis of the dynamics of thematic information flows that are generated in the web space is today one of the most informative methods for studying the relevance of certain thematic areas [69]. This dynamics is due to factors, most of which are not amenable to precise analysis. However, the general nature of the time dependence of the number of thematic publications on the Internet still allows the construction of mathematical models.

Two characteristic features are observed in the behavior of information flows: firstly, an expressive tendency to a constant increase in their volumes, and secondly, the complication of the dynamic structure. Observations of the time dependence of the number of messages in network information flows convincingly indicate that the mechanisms of their generation and distribution are obviously associated with complex nonlinear processes of the overall network dynamics.

Two classes of information flow models are considered traditional: linear and exponential. Both classes have a significant limitation - the monotonic nature of the time dependence, i.e. they are not very suitable for studying the real dynamics of network information flows over long time intervals.

5.2.1. Thematic information flows

Under the thematic information flow we will understand the sequence of messages corresponding to a certain topic. Thus, information systems in our understanding are also thematic information flows, but unlike successive messages in simple information flows, information systems are network structures that take into account numerous information connections.

In a narrow sense, by thematic information flow we mean the number of documents that in some sense correspond to a given topic. Let's consider the general picture of the dynamics of thematic information flows, limiting ourselves to the mechanisms typical for the dynamic segment of the web space.

Numerous facts indicate that in reality the dynamics of thematic information flows is determined by a complex of internal non-linear mechanisms that only partially correlate with the objective environment. It is obvious that this dynamics cannot in principle be explained by a single factor that is fully responsible for all the variety of observed effects. It is this circumstance that explains the great relevance of the problem of modeling the dynamics of thematic information flows.

The information flow, measured by the number of messages, is relatively stable. Only the volumes of message arrays corresponding to one or another topic, one or another information system change in time. In other words, an increase in the number of publications on one topic with a limited ability to generate them (which is quite true) is accompanied by a decrease in publications on other topics [69], so that for each period of time we have T :

$$\int_0^T \sum_{i=1}^{M} n_i(t)dt = NT ,$$

where $n_i(t)$ is the number of publications per unit of time on the topic i; M is the total number of all possible topics, i.e. for local time intervals, one can observe the so-called thematic balance [69].

The main interest in this formulation is the study of the dynamics of a separate thematic flow, which is described by the density $n_i(t)$.

Theoretically, it can be assumed that the sets of publications associated with a certain set of topics overlap, i.e., there are publications that can be simultaneously assigned to several different topics. In reality, such polythematicity is indeed observed, it is an effect that must be taken into account, but in the first approximation, we will assume that its contribution does not distort the overall picture.

Each topic also has a number of characteristic properties that allow some classification, for example, based on the features of its formation and reproduction over time:

- publications on a "one-time" topic, the time dependence of the number of which increases sharply, reaches saturation, and then decreases and then asymptotically tends to zero;

— publications on topics that periodically appear in the general information flow, and then after a while practically disappear from it;

- publications on the topic, the time dependence of the number of which fluctuates around a certain value and never completely disappears.

Thus, messages can be divided into similar categories, and each of them has its own specific development in time.

Even more difficult is a synchronous change in the number of messages from several thematic information flows. Their behavior clearly resembles the processes of interaction between populations in a biocenosis. Thus, in a number of cases, an increase in the number of publications on one topic is accompanied by a decrease in the number of publications on other topics. The general dynamics in this case can be described by a system of equations, each of which refers to a separate monothematic flow. We emphasize that the general polythematic flows are stationary in terms of the number of publications, while the dynamics is mainly determined by the “competitive struggle” of individual topics.

At the same time, in practical terms, a simplified understanding of the information flow as some time-dependent quantity is often completely satisfactory $n(t)$, which is described by the equation

$$\frac{dn(t)}{dt} = F\left(n(t), t\right).$$

Numerous literary sources describe many varieties of "competition" systems for different modifications of the model, depending on a number of assumptions about the actual conditions of the processes. In its simplest form, such equations can take the following form:

$$\frac{dm_i(t)}{dt} = p_i \cdot m_i(t) - \sum_{j=1}^{N_m} r_{ij} \cdot m_i(t) \cdot m_j(t),$$

where N_m is the number of topics.

The above system of equations describes the redistribution of publications between topics that form a fixed set. But
in real life, topics (plots) appear and disappear over time, so it is necessary to introduce appropriate adjustments into these equations. This can be done in different ways, for example, by determining the dependence of the coefficients p_i and r_{ij} on time so that each plot has its own maximum activity in a certain period of time.

5.2.2. Traditional Information Flow Models

Linear model

In some cases, the dynamics of thematic information flows, expressed by the number of publications for a certain period, their intensity, due, for example, to a change in the activity of the topic (its increase or aging), is linear, i.e. the number of messages at time *t* can be represented by the formula

$$y(t) = y(t_0) + v(t - t_0),$$

where t_0 is the starting time of counting ; $y(t)$ *is* the number of messages by the time t; v — the average rate of increase (decrease) in the intensity of the thematic information flow.

Important characteristics of the information flow can be quantified by the fluctuation of this flow - the change in the standard deviation $\sigma(t)$, calculated as follows:

$$\sigma(t_n) = \sqrt{\frac{1}{n}\sum_{i=0}^{n}\left[y(t_i) - \left(y(t_0) + v(t_i - t_0)\right)\right]^2}\ .$$

If this value changes in proportion to the square root of time, then the process of changing the number of publications on a chosen topic can be considered a process with independent increments. At the same time, links with previous thematic publications can be neglected.

In the case when the standard deviation is proportional to some degree of time: $\sigma(t) \propto t^{\mu} \quad (1/2 \le \mu \le 1)$, then the greater the value of μ, the higher the correlation between current and previous messages in the information flow.

Exponential Model

In some cases, the process of changing the relevance of the topic (increasing or decreasing the number of thematic messages in the information flow per unit of time) is approximated by an exponential dependence, which is expressed by the formula

$$y(t) = y(t_0)\exp\left[\lambda(t - t_0)\right],$$

where λ is the average relative change in the intensity of the thematic information flow.

In reality, the relevance of the topic is a discrete value, measured at points in time $t_0,...,t_n$, which is only approximated by the above dependence. Within the framework of this model, the relation

$$y(t_i) = y(t_0)\exp\left[\lambda(t_i - t_0)\right] =$$

$$= y(t_0)\exp\left[\lambda(t_i - t_{i-1} + t_{i-1} - t_0)\right] = y(t_{i-1})\exp\left[\lambda(t_i - t_{i-1})\right].$$

Where

$$\frac{y(t_i)}{y(t_{i-1})} = \exp\left[\lambda(t_i - t_{i-1})\right].$$

Let's introduce the notation: $\lambda(t_i)$ - relative change in the intensity of the thematic information flow at the moment of time t_i :

$$\lambda(t_i) = \lambda(t_i - t_{i-1}).$$

Taking the logarithm of the above equation, we obtain

$$\lambda(t_i) = \ln \frac{y(t_i)}{y(t_{i-1})}.$$

The relative change in intensity at a point in time t_i is also often calculated in practice as the ratio

$$\lambda(t_i) = \ln \frac{y(t_i)}{y(t_{i-1})} \approx \frac{y(t_i) - y(t_{i-1})}{y(t_{i-1})}.$$

The change in fluctuations of a value $\lambda(t_i)$ relative to the mean value can be estimated from the standard deviation:

$$\sigma(t_n) = \sqrt{\frac{1}{n}\sum_{i=0}^{n}(\lambda(t_i) - \lambda)^2}.$$

In this case, if $\sigma(t)$ it changes in proportion to the square root of time, then we can also speak of a process with independent increments - the correlation between individual messages is insignificant. In the case of a significant dependence of messages, the relation is fulfilled: $\sigma(t) \propto t^{\mu}$, where the value μ is greater than 1/2, but limited to 1.

Value μ, which exceeds 1/2 indicates the presence of long-term memory in the information flow. Such a class of processes is called self-similar, for which a correlation is provided between the number of messages published at different points in time.

Logistics model

processes of both increase and decrease in the number of documents take place. Therefore, to build a realistic picture, of course, it is necessary to apply more flexible models.

First of all, it is worth saying that documents in the information flow in many respects resemble populations of living organisms. They are in a certain sense "born", "die" and give birth to "offspring" (documents containing information that has previously appeared in other documents). In modern scientific literature, the concept of a population is often used in a broad sense, and therefore its introduction in modeling information flows is fully justified.

In the second half of the 20th century, significant progress was made in the construction of various mathematical models of population dynamics, in particular, the logistic model, which turned out to be applicable in many branches of science and technology.

The logistic model [66] can be considered as a generalization of the Malthus exponential model, which provides for the proportionality of the rate of increase of the function $y(t)$ at each moment of time to its value:

$$\frac{dy(t)}{d(t)} = ky(t),$$

where k **is** some coefficient.

In real life, as a rule, dynamic systems have sufficiently effective feedbacks that allow you to correct the nature of the processes occurring in them, and thereby keep them within certain limits. Information operations, correcting these feedbacks at certain periods of the evolutionary process, can effectively influence the behavior of the entire system.

The simplest generalization of the Malthus law, which allows you to get away from the unlimited increase of the solution, is the replacement of a constant coefficient k by some function of time $k(t)$. Naturally, this function must be chosen in such a way that the following conditions are met:

— the solution of the equation would have an acceptable behavior;

- the structure of the function would have a certain meaning from the point of view of the phenomenon under study.

The main idea of the logistic model is that in order to limit the growth rate, $y(t)$ an additional condition is imposed on the function, according to which its value should not exceed a certain value. To do this, choose $k(t)$ this type:

$$k(t) = k \cdot [N - ry(t)],$$

where N is the limit value that the function $y(t)$ cannot exceed ; r is a coefficient that describes processes that are negative for a given trend ; k - coefficient of proportionality. Moreover, it is assumed that always $n_0 \le N$. Then instead of the first equation we have

$$\begin{cases} \dfrac{dy(t)}{d(t)} = ky(t)(N - ry(t)), \\ y(t_0) = y_0. \end{cases}$$

A model based on the above equation is called a logistic model. Although ostensibly simple, such a generalization of Malthus's law is by no means primitive. On the contrary, it makes it possible to explicitly include extremely important feedback in the description of population dynamics. The logistic equation can be considered phenomenological: researchers do not need to know how specific mechanisms operate, which, as they increase, $y(t)$ reduce the rate of its change.

There are two classes of solutions to the logistic equation, which, depending on the values of the coefficients, describe the increase and decrease $y(t)$. Their typical behavior is shown in Fig. 23.

The above logistic equation has two equilibrium solutions: $y(t) = 0$ and $y(t) = N$. From a formal point of view, the first of them is unstable, but in practice this is not entirely true. The fact is that the real volumes of information flows are expressed in discrete numbers, and if at some point $y(t)$ it takes on a value less than one, then it will not be able to increase in the future. Therefore, in reality, the solution $y(t) = 0$ can also be considered an equilibrium one.

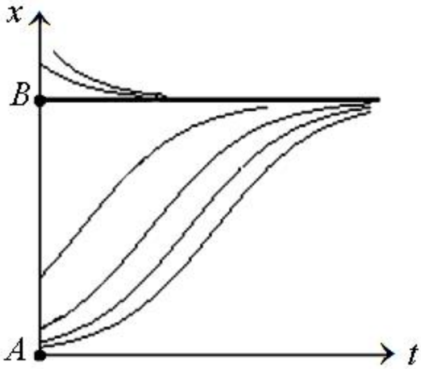


Figure 23. Generalized logistic model

The second solution is $y(t)=N$ equilibrium in any sense. Indeed , $y(t)>N$ at $y(t)<N$, respectively, increasing.

Let us consider the behavior of the dynamics of the thematic information flow, the volume of which is determined by the logistic equation. On fig. 23 shows the resulting dependence of the information flow volumes on time under different initial conditions. At points A and B the rate of change in the number of messages tends to zero: these are stationary states. Between A And B the speed is positive (the number of messages increases), and above the point B it is negative (the number of messages decreases).

The model assumes that subsequently a stationary regime will be established B , which looks quite natural: a larger information flow decreases, and a smaller one increases.

The logistic model satisfactorily describes numerous saturation phenomena. Near A , when the amount of information flow is small, it is very close to the Malthusian model. But for sufficiently large, x there is a sharp difference from the Malthusian growth: instead of tending x to infinity, the number of publications approaches a stationary value B .

Consider how the logistic model can be applied during the analysis of information flows, namely, we will determine the minimum initial number c of messages (which can, for example, be allocated to start some information operation). Let x be the volume of the thematic information flow. The dynamics of this value is influenced by other topics that reduce its distribution, which is described as follows: $\dot{x}=x-x^2-c.$

Calculations show that the behavior of the system changes dramatically at a certain critical value c (Fig. 24).

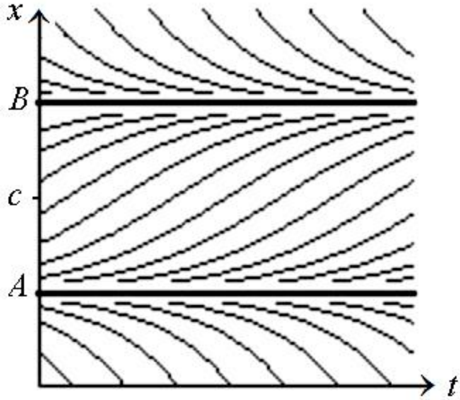


Figure 24. Two fixed points of the logistic model

The system has two equilibrium states A and B. The state B is stable: the amount of information flow in this case is slightly less than in a non-competitive situation. This volume is renewed at small deviations from the equilibrium value B.

The state A is unstable: if for any reason the volume of the information flow drops below the level A, then in the future the number of thematic messages will be reduced to nothing in a quite finite time.

Obviously, in the presence of favorable external conditions (at a certain density of the resource), the volume of the information flow increases freely, which contributes to the logistical age -

niyu. In this case, even more complex models should give results similar to those given. At the same time, this means that the main parameters for specifying the general model can be determined
as a result of the analysis of a simplified logistic model.

Consequently, the logistic model successfully describes the achievement of a certain equilibrium state by the thematic information flow.

In the general case, information dynamics can be represented as a process caused by the emergence and disappearance of individual topics that take place against the background of general trends in the information space. Let's fix a certain topic and assume that $t = 0$ there are n_0 background publications at the moment of time. Due to the fact that (within the framework of the adopted model) the relevance of the topic is maintained for a period of time λ, we can consider separately two time domains: $0 < t \le \lambda$ c $D > 0$ and $t > \lambda$ c $D = 0$ (within the framework of this model, $D = \text{const}$ for each domain, the level of relevance of the topic) and, accordingly, the functions $u(t)$ and $v(t)$, which are the solutions for these regions and are “joined” at the point λ:

$$y(t) = \begin{cases} u(t), & 0 < t < \lambda, \\ v(t), & t > \lambda, \\ u(t) = v(t), & t = \lambda. \end{cases}$$

The first area corresponds to the process of increasing the number of publications in conditions of non-zero relevance of the topic and, possibly, the transition to a state of saturation.

The reaction of the media is never instant: there is always a certain delay in time. This aspect is taken into account in the model by introducing a delay factor τ .

The corresponding dynamics is described by an equation, which, after redefining the coefficients and normalizing them for the function, $u(t)$ can be represented as

$$\frac{du(t-\tau)}{dt} = pu(t-\tau)(1 - qu(t-\tau)) + Du(t-\tau),$$
$$u(0) = n_0.$$

We emphasize that meaningfully the value p determines the normalized probability of the appearance of a publication per unit of time, regardless of the relevance of the topic. This factor reflects the background mechanisms of information generation (a typical example would be mechanical reprinting of materials from prestigious information sources). The value D characterizes the direct impact of the relevance of this topic. The parameter q characterizes the decrease in the growth rate of the number of publications and is the inverse value of the asymptotic value of the dependence $u(t)$ at $D=0$.

For the second region described by the function $v(t)$, respectively, we have

$$\frac{dv(t-\lambda)}{d(t)} = pv(t-\lambda)(1-qv(t-\lambda)).$$

In this case, the condition of equality of functions $u(t)$ and $v(t)$ at the moment should be taken into account $t=\lambda$:

$$v(\lambda)=u(\lambda).$$

The above non-linear differential equations are variants of the Bernoulli equation:

$$y' = ay^2 + by,$$

which is linearized by the standard change $z=1/y$:

$$z' + bz + a = 0.$$

The general solution of this equation has the form

$$z = \frac{1}{\mu(x)}\left[C - a\int \mu(x)dx\right]$$

with integrating factor

$$\mu(x) = e^{bx}.$$

Variables C are defined as follows: for the first region from the initial conditions, and for the second - from the "crosslinking" condition. As a result of simple transformations, we find a solution for the first region:

$$u(t) = \frac{u_s}{1 + (u_s / n_0 - 1)\exp\left[-(p+d)(t-\tau)\right]},$$

where u_s is the asymptotic value u, the value of which determines the saturation region:

$$u_s = \frac{p+D}{pq}.$$

Thus, the model describes a dependence that has *an S*- like (logistic) form (Fig. 25).

Note that the solution does not depend on the value of n_0, which indicates the insignificance of the initial conditions for information dynamics. Whatever the initial number of publications, saturation will be determined solely by parameters characterizing the background rate of increase in the number of publications, a quantitative measure of relevance, and factors negative for the process.

The curve presented in fig. 25 has an inflection point

$$t_{\inf} = \frac{1}{p+D}\ln(u_s / n_0 - 1) + \tau.$$

Thus, for the first region we have the so-called *S*- like dependence, and for $t \sim t_{\inf}$ behavior $u(t)$ approaches _

linear dependence and corresponds to the linear model.

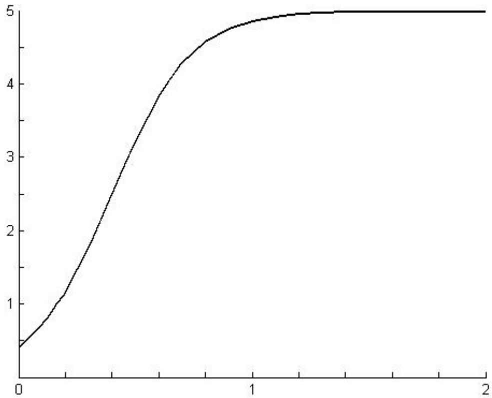


Figure 25. Increasing function

Let us now present the expression for $u(t)$ in the following form:

$$u(t) = \frac{u_s}{\exp\left[(p+D)t\right] + (u_s / n_0 - 1)\exp\left[(p+D)\tau\right]},$$

from which it can be seen that under the condition

$$t < \frac{1}{p+D}\ln(u_s / n_0 - 1) + \tau = t_{\text{inf}}$$

the dependence $u(t)$ has an exponential character, i.e. for values of *t* significantly smaller than t_{inf} , the model coincides with the exponential model.

For the second region, respectively, we have (Fig. 26)

$$v(t) = \frac{v(\lambda)}{qv(\lambda) + (1 - qv(\lambda))\exp\left[-p(t-\lambda)\right]^2},$$

taking into account the condition of "crosslinking": $v(\lambda) = u(\lambda)$.

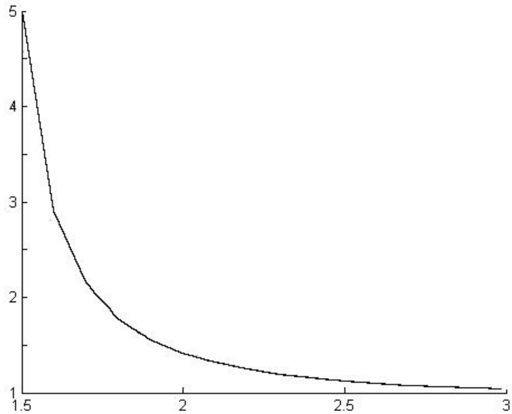


Figure 26. Decrease function

If the dependence $u(t)$ manages to reach saturation in the time interval $t < \lambda$, then the above equation can be simplified by presenting it as follows:

$$v(t) = \frac{v_s(p + D)}{p + D(1 - \exp[-p(t - \lambda)])},$$

where $v_s = 1/q$ is the asymptotic value of the dependence $v(t)$.

As expected, the value v_s also does not depend on either the initial condition or the condition of "matching" with the function $u(t)$ on the boundary of the regions. As can be seen, the resulting dependence has a saturation region u_s at $t \le \lambda$ and asymptotics v_s, which describes a gradual decrease in the number of publications to the background level, i.e. it, at least on a qualitative level, agrees with the general considerations about the nature of information dynamics obtained on the basis of experimental data. In addition, in local areas it is well approximated by linear and exponential models. A typical full dependence $y(t)$ is shown in fig. 27.

In the case of information flows that are associated with specific topics, it is necessary to describe the dynamics of each of these flows separately, taking into account that an increase in one of them automatically leads to a decrease in others and vice versa. Therefore, the limit on the number of messages on all topics also applies to the totality of all monothematic streams. In the case of studying the general information flow, there is a phenomenon of “flowing” of publications from some topics that are losing relevance to others.

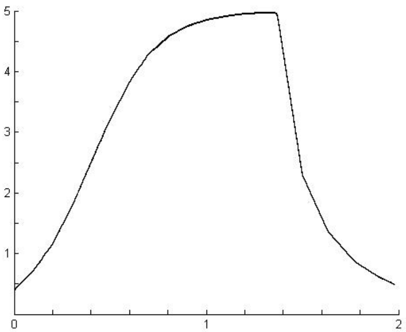


Figure 27. Generalized graph of dynamics thematic information flow

The general dynamics should be described by a system of equations, each of which belongs to a separate monothematic flow. We emphasize that the general polythematic flows are stationary in terms of the number of publications, while the dynamics is mainly determined by the “competitive struggle” of individual topics.

The above system of equations of "competition" in the framework of a generalized logistic model can be represented as follows:

$$\frac{dy_i(t)}{dt} = \left(p_i + D_i(t, \lambda_i) \right) \cdot \left(y_i(t) - \sum_j r_{ij} \cdot y_i(t) y_j(t) \right).$$

In these relations, the coefficients p_i and D_i have the same meaning as before, and λ_i are the points at which the corresponding D_i reach maximum values.

5.2.3. Identification of information clusters

In practice, when searching for news information, the task always arises of identifying information clusters (plots) consisting of separate documents and ranking them according to some criteria, which should ensure not only the identification of the most important topic, but also a “fan-shaped” multi-aspect coverage of all the most significant aspects. infoplots. This problem is solvable in many systems using various approaches and algorithms. At the same time, the technological chain remains unchanged: building a semantic network from information messages, clustering - identifying the most interconnected groups, i.e. infoplots, "weighing" (assessment of importance, relevance) and visualization of the most significant of them [66].

When highlighting story chains to determine the pairwise textual proximity of individual documents, as a rule, algorithms for identifying such documents are used, which have already become traditional in search engines. Thus, the pairwise proximity matrix of documents is processed by clustering algorithms such as *LSA/LSI* , *k*- means, suffix trees, etc. The selected classes of documents represent information plots.

For presentation to users, information stories must be ranked. The main factors influencing the ranking by importance are the speed of information and the size of the story chain. Efficiency is understood as some function of the time of publication of all documents in the information story, and the size reflects the general interest in a particular topic. In all these approaches, the central task is to identify documents related to the same story and to identify "non-overlapping" stories. On fig. 28 shows a typical approach to identifying infoplots. The last information message (document) generated by the generation is compared with the previous ones , the level of their similarity is estimated. If the similarity level exceeds a certain threshold, then the analyzed document is considered to belong to the information plot to which the previously generated document belongs. If there are no such documents, a new plot is fixed, which currently consists of one document.

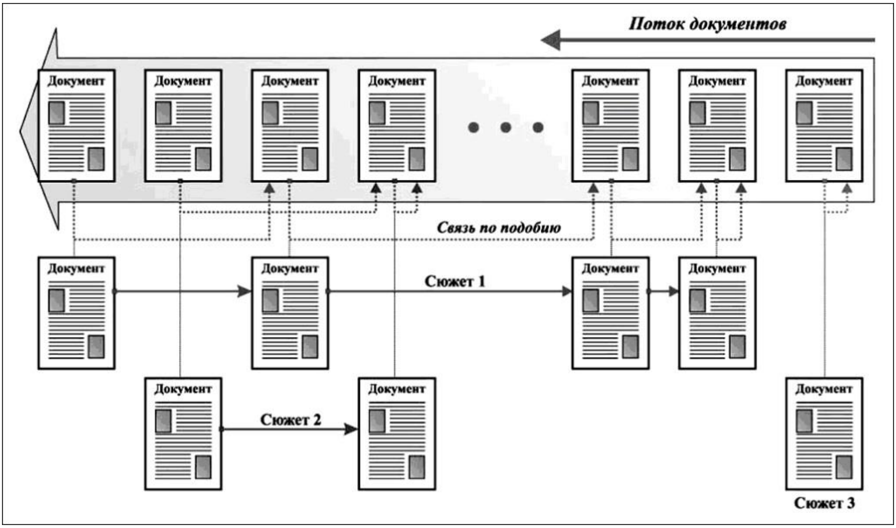


Figure 28. Typical scheme for identifying news stories

For the resulting display of each individual infoplot, documents selected by textual proximity from various sources, sorted in chronological order, are used. At the same time, stories can be digests integrating common places of documents on the topic, as well as unique information contained in individual documents. In other words, summarizing the plot in this case does not come down to collapsing information, but to building an extended version, in comparison with any document from the plot chain.

For example, in the Yandex.News system (http://news.yandex.ru), to highlight an information plot, a matrix of pairwise proximity of documents is constructed, which is processed by a clustering algorithm with empirically selected parameters (in particular, the radius of the proximity metric). In order to increase the connectivity of large plots, second-level clustering is additionally used, which ensures the collection of atomic clusters into larger ones. All messages in the search results are grouped on the site, while the ranking of infoplots is based on the Yandex standard ranking principles. It is based on the number and rank of news items within news stories, with the rank of one news item being defined as its "freshness", taking into account the priorities for satisfying the search criteria.

In the InfoStream system (http://infostream.ua), the thematic similarity of documents is determined on the basis of normalized sequences of the most significant terms included in each document [70].

Sequences of similar (with a certain coefficient of mutual proximity exceeding some empirically established level) documents form chains. In this case, each document falls into some chain, in extreme cases, consisting only of itself. Then the chains are evaluated in terms of length and efficiency, after which the user is presented with a certain number of the most important thematic news stories. To represent the story chain, the document titles are also evaluated relative to the keywords corresponding to the plot, and then the most "weighty" titles are selected from all the titles for display (Fig. 29).

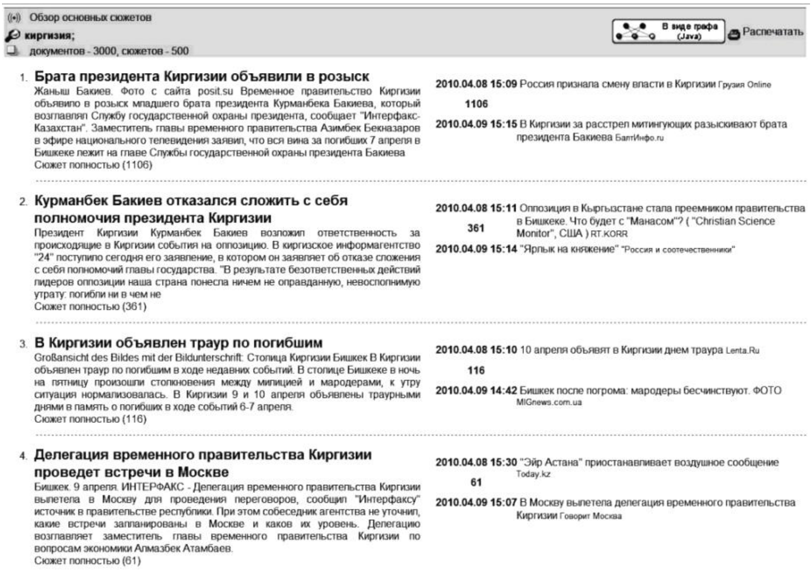


Figure 29. An example of displaying infoplots in the InfoStream system

5.2.4. Emergence of information systems

In complex systems (and modern information systems, no doubt, they are), among many other characteristics, integrity is most clearly manifested, i.e. the presence of such properties that are not inherent in any element (in this case, the document) that makes up the system, taken separately. The property that is called "emergence" is the result of the emergence of special synergistic relationships between the elements of the system. The term “emergence” introduced by F. Lewis means that in systems the whole is more often than the sum of parts [71], i.e. at each level of complexity, new, often unforeseen qualities arise that are not inherent in the constituent parts.

So, if we consider a clock as a system - a device that shows the current time, then none of its details will be able to show time. It can't even show "part of the time". The ability to show time appears in all parts together, and after they are in a certain way assembled into a single complex and, thereby, enter into a certain interaction with each other.

The emergence of an information system makes it possible not to be limited to the study of its elements and the relationships between them, but to perform a holistic analysis of the entire system. Until the end of the 20th century, in the analysis of complex systems, including social systems, the reductionist approach was mainly used, which explained many properties of complex systems by the properties of their elements - "atoms" or "molecules". As a result of the development of system analysis, the emergence of the science of the complexity of a technological breakthrough in computing capabilities, the situation has changed dramatically. At the moment, such areas as the theory of chaos, complex networks, nonlinear and self-organizing systems have been developed. It turned out that many properties of complex systems cannot be derived from a predetermined set of dynamic equations. On the contrary, the equations can only be obtained as a result of numerical simulations.

At the same time , it is obvious that it is impossible to develop and apply in practice some universal methodology for modeling information systems. This is mainly due to the weak formalization of many concepts and factors, primarily subjective. In each individual case, one has to trust the knowledge and intuition of analysts who are professionally involved in the analysis of information processes.

With objective factors, the situation is different. They are fully amenable to analysis at the statistical level and allow quantitative estimates that can be used to make reasonable forecasts. Modern methods of applied statistics, time series analysis include a large arsenal of detailed and tested methods. However, statistics makes it possible to describe only the formal aspects of the phenomena being studied, leaving the substantive aspects “overboard”. Therefore, there is a need to expand the set of tools that are used in the analysis and modeling of information flows. One of the most promising areas in this regard is mathematical modeling. Today, mathematical modeling is widely used in many branches of science and technology, while the modeling of information systems remains an open problem.

In the field of information systems, modeling is promising, due to some realistic rules for the behavior of individual elements (documents, topics), which are refined by some parameters that are changed during modeling. In this case, the inverse problem also acquires great value - to estimate the value of the model parameters from the real behavior of some dependence. Knowledge of the general behavior of sustainable solutions makes it possible to predict the development of information systems even when there is no exact understanding of the specific mechanisms that determine their dynamics, and such predictions may turn out to be more accurate than those obtained by traditional expert methods. If the decisions turn out to be unstable, then important information about the system can also be obtained from this, allowing in some cases to predict in which direction the dynamics of individual information systems is directed.

Attempts to model information flows have been made for a long time, but they were hampered by computational difficulties, especially when it was necessary to describe the dynamics of feedback systems. Today, there are a sufficient number of possibilities for efficient computer data processing, which allows, on the one hand, to prepare sets of input pairs - meters based on the analysis of the results of statistical studies, and on the other hand, to solve formalized problems with a sufficient degree of accuracy and within an acceptable time. All this gives reason to believe that in the near future mathematical modeling will become the main tool for analyzing and managing information flows.

One application of the concept of emergence to modeling today is multi-agent modeling (see Chapter 2). Multi-agent models are widely used to analyze decentralized systems, the regularities of the functioning dynamics of which have not been sufficiently studied. These models are used to study the general behavior of complex systems, to identify the rules for their functioning, taking into account assumptions about the individual behavior of its individual components.

5.2.5. Synergistic approach

Under the influence of the external environment, information systems can move to unpredictable behavior - chaos.

Disordered, unpredictable, random behavior of the system is associated with non-deterministic chaos, in which it is impossible to derive patterns for determining the future state of the system, knowing its previous state. Today, more and more attention of scientists is drawn to deterministic chaos, which is generated not by the random behavior of a large number of system elements, but by the internal essence of nonlinear processes. The behavior of information systems fully corresponds to the definition of deterministic chaos. For complex systems, which, of course, are information systems, the equations describing their behavior turn out to be so complex that they cannot be solved by analytical methods. Therefore, their study is usually carried out by means of computer simulation.

When solving nonlinear problems, the state of the system and the degree of its organization are depicted using the so-called phase space, the coordinates in which are the parameters characterizing the system. For example, to describe systems in mechanics as phase space coordinates, the positions of individual points and their velocities are used. In this case, deterministic chaos is displayed as a continuous trajectory, which sometimes can gradually fill the entire phase space (any small neighborhood of a point in the phase space will be crossed by many phase trajectories). This property of deterministic chaos leads to the concept of fractals, the fractal dimension, for example, the Hausdorff dimension of a trajectory densely covering a plane cannot be an integer. Fractional dimension is one of the main features of fractals.

The key concepts of synergetics are "bifurcations" and "attractors". The bifurcation point is usually understood as the state of the system, after which a certain set of options for its development is permissible. That trajectory, or that set of trajectories, along which the development of the system after the bifurcation point is possible, and which differ from others in relative stability, are called attractors. In other words, the attractor seems to attract to itself a set of trajectories possible after the bifurcation point. The properties of bifurcation points and attractors are studied in the theory of complex systems, where the laws of development of such systems, transitions from chaos to order and vice versa are established.

Indeed, bifurcation can lead to chaos. Let us describe an abstract but rather convincing example, M. Feigenbaum's cascade of bifurcations, one of the typical scenarios for the transition from a simple periodic regime to a complex aperiodic regime with an infinite doubling of the period [72]. The Feigenbaum sequence has a self-similar, fractal structure - an increase in some area demonstrates the similarity of the selected area with the entire structure.

Feigenbaum analyzed mainly the logistic equation $X_{n+1} = CX_n - CX_n^2$ (C is an external parameter), from which he deduced that under certain restrictions in all such equations there is a transition from an equilibrium state to chaos.

The logistic equation, which, as is known, has two stable solutions, is usually treated as a condition of population dynamics and allows the following interpretation: it is assumed that there is a population of individuals with a specific number X_n in isolation. A year later, offspring appear with a specific number X_{n+1}. The population growth is described by the first member of the right side of the equation (CX_n), where the coefficient C determines the growth rate and is the determining parameter. The decrease in the population size (due to overpopulation, lack of food, etc.) is determined by the non-linear term (CX_n^2). The calculation results (Fig. 30) show that:

- at $C < 1$ the population dies out with an increase n;

— in the region, $1 < C < 3$ the population size approaches a constant value $X_0 = 1 - 1/C$, which is the region of stationary, fixed solutions. With a value, $C = 3$ the bifurcation point becomes a repulsive fixed point;

- bifurcations begin to appear in the range $3 < C < 3,57$ - a branching of each curve into two (indeed, the logistic equation has two stable classes of solutions). The population size fluctuates between two values that lie in these areas. At first the population increases sharply, the following year there is overcrowding, and a year later the number decreases again;

- at $C > 3,57$ there is an overlap of different areas, and the behavior of the system becomes seemingly chaotic.

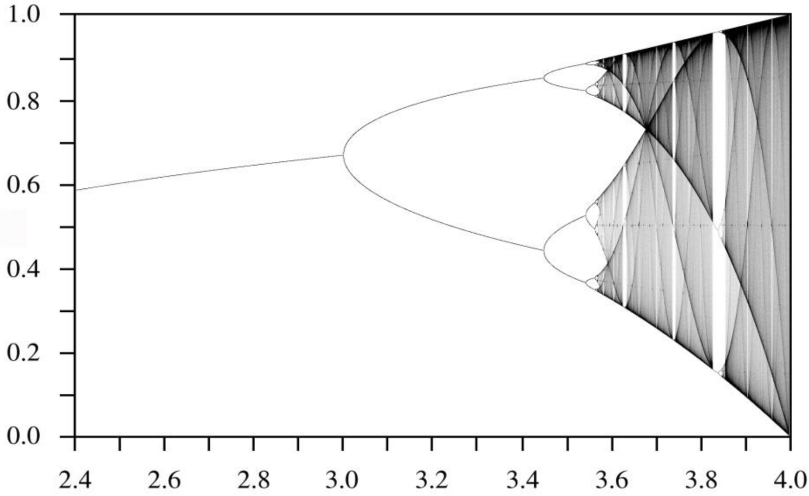


Figure 30. Cascade of bifurcations (Fegenbaum sequence): the abscissa axis is the value of the C parameter , the ordinate axis is the value X_n

Thus, the final state of evolving physical systems is the state of dynamic chaos.

Using the theory of bifurcations, it is possible to predict the nature of the behavior that manifests itself during the transition of the system to a qualitatively different state, as well as the area of existence of the system and evaluate its stability.

Bifurcations very often arise during the transition of a system from a state of apparent stability and equilibrium to chaos. On the basis of chaos, it is not only impossible to build or refine a forecast, but also, accordingly, to check it. However, this should not indicate the inaccuracy of the theory of chaos, confirmed both in mathematical calculations and in life. At present, there is no mathematically exact application of chaos theory for the study of information flows, however, this theory is already able to provide for the transition of system models presented in an analytical form into a chaotic state. Obviously, this is indeed one of the most promising areas of applied research of information processes.

In synergetics, it is strictly proved that it is impossible to “impose” the desired behavior on the system by any external influences - you can only choose the best possible trajectory from the potentially laid ones [73].

Therefore, when planning and modeling information systems, one of the main tasks is to find the bifurcation points of information processes and the formation of fluctuations that determine the choice of the required evolution trajectory (attractor).

As you know, in mathematics, catastrophes are called spasmodic changes that occur as a response of the system to a smooth change in external conditions. Information systems can cause processes that are best described in terms of catastrophe theory: “Near bifurcation points, significant fluctuations can be observed in systems. Such processes seem to hesitate before choosing one of several paths of evolution. A small fluctuation can serve as the beginning of evolution in a completely new direction, which will dramatically change the entire behavior of the macroscopic system” [74]. This explains why it is so difficult to prevent a catastrophe when its signs are already visible: the speed of its approach at this time increases infinitely [75].

If we consider society as a complex system, then information systems can be considered as a means of influencing it, as was shown above for choosing certain development paths. Thus, models of information systems are part of more general social models.

5.2.6. Game-theoretic approach

A characteristic feature of many information systems is the fact that their components (actors) are in a state of conflict of interest, and at the same time they operate in the absence of complete information about each other's intentions. In particular, when analyzing information processes, one almost always has to analyze conflict situations in which the interests of two or more competing parties pursuing different goals collide. Mathematical theory, which is devoted to the study of conflict situations, is the theory of games. Thus, it seems quite natural to try to apply game theory to the study of information systems. In a generalized game (for example, a set of information operations of opposing forces can be considered as a game), the interests of two or more opponents may collide. In this case, players can form coalitions, in which case the game is coalitional.

The structure of any game is described by three blocks:

1) admissible sets of moves or strategies of participants;
2) the goals of the participants;
3) the type of behavior and awareness of the participants.

In game theory, games are classified as cooperative (cooperative) and non-cooperative.

In cooperative games, participants can form groups, taking on some obligations to other players and coordinating their actions. In this they differ from non-cooperative games in which everyone is obliged to play for themselves.

Of the two types of games, non-cooperative ones describe situations in great detail and produce more accurate results. Cooperatives consider the process of the game as a whole.

Below we will focus on non-cooperative games. Formally, a troika is called a kind of operational game $\Gamma = \langle I, X_i, H_i \rangle$, where I is the set of participants in the game; X_i is the set of participant strategies , $i \in I$; H_i is the participant's payoff function, $i \in I$, defined on the set of situations (concrete implementations of the strategies of all participants in the game), mapping it to the set of real numbers.

The non-cooperative game assumes the following order of play:

1. Players, simultaneously and independently of each other, choose their strategies from the set. X_i The strategy vector $x = (x_1, x_2, ..., x_n)$ of all players represents the situation in the game.

2. Each participant receives a payoff determined by the value of the function $H_i(x)$; this stops the interaction between the players.

Game analysis is the ability to predict the solution of the game - the set of possible moves and their results. Important concepts in game theory are also the optimal strategy, the price of the game, and the average payoff. In particular, the strategies P^* of the first and Q^* second players are called optimal, and the number V is called the cost of the game, if the following inequalities hold for any strategies P of the first and Q second players [76]:

$$M(P, Q^*) \le V \le M(P^*, Q),$$

where is $M(P, Q) -$ the mathematical expectation of the payoff of the first player who chooses the strategy P , provided that the second player chooses the strategy Q .

In many problems from game theory, uncertainty is caused not by the opposition of the opponent, but by the lack of awareness of the player about the conditions in which the parties operate, for example, about information operations against him, information influences. Such games are usually called "games with nature", in the solution of which the so-called risk matrices are used. Within the framework of this approach, the actions of an intelligent *ratoinal decision-maker* are determined by its awareness of the state of the environment and the views of opponents. In this case, information about the environment transmitted to the agent, as well as about the ideas of opponents, can be an element of information impact.

The information transmitted to the agent for informational impact can be [77]:

- "dry" facts;
- logical conclusions, analytical judgments based on a certain set of facts;
- Emotionally colored statements.

As data, a forecast can also be transmitted to the agent, depending on an indefinite parameter and the actions of the agent itself. Each agent, on the basis of an "active forecast", can "restore" information about the environment and use it when making decisions (for example, when calculating equilibrium actions).

When solving problems under conditions of uncertainty, when the probabilities of individual particular outcomes are unknown, difficulties arise in mathematical modeling. Therefore, decision theory, in particular, recommends applying an approach based on the well-known Bayes theorem. The optimization strategy in such cases is based on the Bayesian decision theory. In this case, the loss function accepted in game theory is considered as a generalization of the error probability. Accordingly, it is supposed to choose a solution that minimizes the loss function. The Bayesian approach to assessing probabilistic relationships plays a crucial role in the theory of decision making under conditions of uncertainty in the consequences of these decisions or in the face of opposition from nature or competition. Under these conditions, the key is the control strategy based on the a posteriori (postexperimental) probability of the event. A prerequisite for the correctness of this approach is the constant training of the system. The control strategy is first built on the basis of certain ideas about the probability of events, and as the system functions, the control correction is implemented - the use of accumulated experience by recalculating strategy options taking into account the changed probabilities.

Note that the application of game theory has two different aspects: firstly, it can be used to optimize the decision-making mechanisms of the opposing sides, and secondly, to develop the principles of their organization. In particular, in the second case, the question of stability becomes extremely relevant.
game (by which we describe the electoral process) in the sense of Nash. Formally, the Nash equilibrium is defined as follows [78]. In cases where each player builds his expectations about the partner's behavior based on the past experience of similar games, a game solution that is stable in some sense is called the equilibrium of this population. Then the Nash equilibrium becomes of particular importance - a profile of strategies from which it is not profitable for anyone to deviate if the partners do not deviate, i.e. A game is said to be stable in the sense of Nash if neither player can increase his payoff solely as a result of his own actions. Nash equilibrium (*NE) is a point from which* it is not profitable for any player to leave during the current moves of partners, and strict Nash equilibrium (*SNE)* is a point from which it is unprofitable to leave. When each player $i \in I$ chooses a strategy x_i from the strategy vector $x = (x_1, x_2, ..., x_n)$, the player i receives a payoff $H_i(x)$. At the same time, the payoff of the i-th participant in the game depends on the entire profile of strategies: not only on the strategy chosen by the player himself i, but also on other people's strategies. The vector of strategies $x*$ is a Nash equilibrium if no player benefits from changing his strategy, i.e. for any i condition is true:

$$H_i(x*) \geq H_i(x_i, x_{-i}*).$$

Here $x_i, x_{-i}*$, is a vector composed of all coordinates of the vector $x*$ except i-th , which corresponds to the value x_i.

A game can have a Nash equilibrium in pure strategies or in mixed strategies (i.e. choosing a pure strategy stochastically at a fixed frequency). Nash proved that if mixed strategies are allowed , then in every game *of n players there will* be at least one Nash equilibrium.

When analyzing social processes, situations of asymmetric conditions for different players are often considered. In such cases, it makes sense to consider the Stackelberg equilibrium [79], which, in contrast to symmetric conditions, implies different principles for the formation of expectations of different players.

The first player (leader) is guided by the optimal responses of partners, knowing their preferences, and the rest play, as in the case of the Nash equilibrium, only reacting to his move and to each other's moves. The Stackelberg equilibrium can arise, for example, when one of the players makes his choice before the others and knows their goals. Or when he is alone, and there are enough followers of the same type so that each of them cannot calculate the overall consequences of his move.

Consider a special case of the Stackelberg model - the struggle of two information systems for the electoral preferences of people. Let there be two parties, one of which is the "leader" and the other is the "pursuer". Let campaign costs be a linear function of the total electorate Q:

$$P(Q) = a - bQ.$$

We also assume that the costs of each of the two parties (advertising, information systems, local information operations, etc.) per supporter are constant and equal to c_1 and , respectively c_2. Then the conditional "profit" of the first batch will be determined by the formula

$$\Pi_1 = P(Q_1 + Q_2) \times Q_1 - c_1 Q_1,$$

and the number of votes of the second

$$\Pi_2 = P(Q_1 + Q_2) \times Q_2 - c_2 Q_2.$$

In accordance with the Stackelberg model, the first party - the leader - at the first step achieves the number of its supporters Q_1. After that, the second party, the persecutor , analyzes the actions of the leader and achieves the number of supporters Q_2. The goal of both parties is to maximize the conditional profit.

The Nash equilibrium in this game is determined by backward induction. Consider the penultimate stage of the game - the course of the second game. At this stage, the second party knows the optimal number of supporters of the first party Q_1*. Then the problem of determining the optimal number of its supporters Q_2* is reduced to solving the problem of finding the maximum point of the profit function of the second party. Maximizing the function Π_2 with respect to the variable Q_2, for a given Q_1, we find that the optimal number of supporters of the second party is (we assume that $c_1 = c_2 = c$):

$$Q_2^* = \frac{(a - bQ_1^* - c)}{2b}.$$

This is the best response of the pursuing party to the leader's choice of Q_1*. The leading party can maximize its profit function given the form of the function Q_2*. The maximum point of a function Π_1 with respect to a variable Q_1 when substituting Q_2* is defined as follows:

$$Q_1^* = \frac{(a - c)}{2b}.$$

Substituting this into the expression for Q_2*, we get

$$Q_2^* = \frac{(a - c)}{4b}.$$

Thus, at equilibrium, the leading party must acquire twice as many supporters as the pursuing party.

It should be borne in mind that game theory models, to a lesser extent than many of the others considered, can be used for accurate calculations and forecasts. Rather, here we can talk about a well-founded methodology that can significantly increase the effectiveness of the actions of participants in social processes. These models are essentially sets of recommendations that provide measurable benefits to those who use them.

5.2.7. extreme approaches

Extreme approaches to modeling the behavior of complex systems are widely used in the natural sciences, most recently at the intersection of ecology and biology [80], where they are successfully used to study population dynamics—the development of individual populations.

Research conducted in this area, almost without technical changes, can be used to study human communities, social processes, the dynamics of electoral populations, in particular, under the influence of information operations.

In accordance with extreme approaches to modeling, only those states of systems are realized that correspond to the extrema of some objective function (described by equations) under certain boundary conditions. The most subtle issue in this case is the principles of compiling equations, which in the case of the study of information operations (as well as in other areas) are based on the experience of experts, analogies, incomplete empirical patterns.

When modeling information systems, approaches based on the logistic equations of population growth [81], obtained as a result of solving optimization problems, the principles of the stationary state of open systems [82], the principles of maximum population diversity [83], the maximum generalized entropy [84], the maximum of the Malthusian parameter [85] and many others.

Let's take a closer look at some of the existing approaches.

Principle of Survival

When studying the dynamics of information systems, the principle of survival (preservation of the components of the information system and their functional significance) can be used as an optimality criterion, using the mathematical apparatus proposed in [86].

It is assumed that the dynamics of the information system is adequately described by a system of equations, the parameters of which are some social conditions, as well as the structural and functional parameters of all information systems. Allocate s-th information system and some structural or functional parameter α_{s_k} of this system. It can be assumed that the information system consists of two subsystems that differ in the value of some phenotypic parameter (characteristics inherent in the components at a certain stage of development). Let $x_s^{(1)}$, $x_s^{(2)}$. $\alpha_{s_k}^{(1)}$. $\alpha_{s_k}^{(2)}$ — the number of components and the functional parameters of the two subsystems.

Modeling the system, in which appropriate changes have been made, taking into account the differences in this functional parameter for the component s-th population, makes it possible to analyze the asymptotic properties of the size of subsystems. One of the possible behaviors is the displacement by the second subsystem of the first, when the parameter $\alpha_{s_k}^{(1)}$ has an advantage over $\alpha_{s_k}^{(2)}$ in a given information situation, i.e.:

$$\lim_{t\to\infty} x_s^{(1)} > 0, \lim_{t\to\infty} x_s^{(2)} = 0.$$

From the point of view of displacement by the second subsystem, $\alpha_{s_k}^{*}$ the optimal value of the parameter is such $\alpha_{s_k}^{(1)}$ that for any parameter value different from this, the $\alpha_{s_k}^{(2)} \neq \alpha_{s_k}^{*}$ above conditions are satisfied for any initial states of the system. In some cases, the information system may not have such an optimal parameter value, i.e. the system can stably exist at any value of the parameter α_{s_k} , which belongs to the area corresponding to the condition for the stable existence of the information system, even if the value is not equal to the optimal one. The optimal value is established as a result of the competition of individuals with different values of the considered functional parameter. It is due to this competition that the components of systems with non-optimal values of the parameter $\alpha_{s_k} \neq \alpha_{s_k}^{*}$ “leave” the information system [87].

Using the selection criterion, in the case under consideration, it is necessary to take into account the limitations arising from the information patterns of the process. For further modeling, the simplest requirement of the maximum relative rate of increase in the size of the information system can be used as an optimality criterion :

$$k = \frac{d \ln x}{dt} = \max.$$

Such a criterion can be applied to determine the optimal values of structural and functional parameters, if the relative rate of increase in the number of information system components is presented as a function of these parameters.

The principle of maximum surprise

One way to study the dynamics of biological populations is to study "Darwinian systems" that describe the dynamics of natural selection. Consider how it can be used to model information systems. In the work of E.V. Evdokimov [88] presented a method for describing Darwinian systems (DS) according to Eigen. Such systems are open, consisting of units of various types self-copying with a small number of errors, using for their reproduction the free energy of nutrient components coming from outside. In the case of information systems, information impacts, including information operations, can be considered as such an external energy impact. With this approach, the constancy of the total number of system elements (the size of the information system) can be a limitation. To describe the DS, the differential equation is used

$$\dot{x}_i = x_i(A_i Q_i - \Delta_i) + \sum_{j \neq i}^{w} u_{ij} x_j - F_i \ ,$$

which simplifies to the following expression:

$$\dot{y}_i = y_i(\mu_i(s) - D),$$

Where $i, j = 1, 2, ..., w \quad (w = \text{const})$ is the number of populations in the system; $s = \{s^1, s^2, ..., s^m\}$— concentrations of "nutritional components" (volumes of impacts); $\mu_i(s)$ is the specific rate of increase i in the th population; D is the flow rate in the system. Depending on the restrictions imposed, a distinction is made between DWs with a permanent organization, in which the sum $\sum_{i=1}^{w} y_i$ and concentration s are constant, and DWs with a constant flow, characterized by the condition $D = \text{const}$.

To solve the problem of the incompleteness of the above equations and the inaccessibility of information at the microlevel, it is proposed to use the postulate [88], which consists in the fact that “the process of evolution of the DS proceeds in the least unexpected way” (the principle of minimal unexpectedness of the course of evolution).

The surprise function of the evolution of a DS with constant organization is used as the objective function:

$$I(P_i(t)/P_{i0}) = P_i(t)\log(P_i(t)/P_{i0}),$$

where $P_i(t) = P_\mu(\mu = \mu_i, t)$ is the probability that t a population randomly selected at the moment i $(i = 1, 2, ..., w)$ has the Malthusian parameter μ_i, $P_{i0} = P_i(t)|_{t=0} = y_i / \sum_{k=1}^{w} y_{k0}$, the values y_{k0} are set experimentally.

In this case, the variational problem is formulated as follows:

$$\begin{cases} \sum_{i=1}^{w} I(P_i(t)/P_{i0}) \to \min; \\ \delta\left[I(P_i(t)/P_{i0})\right] = 0. \end{cases}$$

The solution was obtained by the method of indefinite Lagrange multipliers:

$$P_i(t) = \frac{P_{i0}e^{\mu_i t}}{\sum_{k=1}^{w} P_{k0}e^{\mu_k t}},$$

moreover, it is proved that it corresponds to the solution of the above system of equations for $\dot{y}_i$. Giving the Lagrange multiplier $\lambda_0 = \log \sum_{i=1}^{w} P_{i0}e^{\mu_i t}$ information sense, one can get Fisher's "basic theorem of natural selection":

$$\frac{d\hat{\mu}}{dt} = \sigma_\mu^2.$$

In addition, it is proved that the multiplier λ_0 is proportional to the "energy consumption" of the population.

Thus, based on the heuristic principle of minimal unexpectedness of the evolutionary process, the results describing the dynamics of selection in Darwinian systems are obtained, which are completely identical to the equations derived from the kinetics of reproduction and competition, and the Lagrange multipliers used to solve the variational problem are quite meaningful and have a predictive value. value.

Malthus parameter maximum principle

Let the information space consist of *w* information systems. The information space can be described by the numbers of its constituent information systems x_i. Let $x = \sum_{i=1}^{w} x_i$ be the total number of information space elements. Let's assume that during some local time interval i - th information system is characterized by the Malthusian parameter $\mu_i(t)$ from the equation $dx_i / dt = \mu_i x_i$.Let $p_i = x_i / x$ - the relative share of i - th information system in the information space. Then the set $p = \{p_1, p_2, ..., p_w\}$ is called the structure of the information space; value $\hat{\mu} = (\mu, p)$ is the average Malthusian parameter $((\mu, p) = \mu_1 p_1 + \mu_2 p_2 + ... + \mu_w p_w)$, and the dynamics of the development of the size of the information space is described by the equation $dx / dt = \hat{\mu} x$.

This approach is based on the principle of maximum average Malthusian parameter, i.e. that the community of interacting information systems evolves in such a way that its average Malthusian parameter always increases, reaching its maximum in stable equilibrium. In [87], the structures are divided into probable and improbable, and the conditions are determined under which, in the process of adaptation, all populations are eliminated from the community, except for one or none of the populations leaves the community (in our case, the information space).

5.3. Nonlinear Dynamic Models

It is known that if the state of the system does not change in time, it is called static, otherwise - dynamic. It is clear that in terms of studying information systems, it is precisely dynamic systems that are of primary interest: first of all, we are interested in the ongoing changes, and static systems do not generate any changes. In turn, dynamical systems are divided into two classes: linear and nonlinear. Linear systems are called systems, the characteristics of which depend on the change in the states of these systems. Conversely, the characteristics of nonlinear systems depend on such changes.

Very often, non-linear social systems manifest themselves, first of all, as a disproportionate response to external influences. It is well known that such systems can endure severe shocks with amazing ease and without consequences, and at the same time instantly "go over the top" from an insignificant event or influence. It is information operations (the tool and object of which are information systems) that can be considered as such an influence or impact.

Models used in relation to non-linear systems are also called nonlinear, since a model is also a system, and, of course, it can be non-linear. The nonlinearity of the model can be formally expressed in the structure of the equations used, and their solution in some cases can be quite feasible, at least in numerical form.

The main emphasis in the construction of non-linear competitive models of information processes is currently placed on analysis of fundamental internal interactions of dynamic systems based on logistic models. Modeling of development dynamics based on differential logistic equations is widely used to model a wide variety of both natural and informational processes.

Naturally, before applying mathematical models, it is necessary to substantiate their adequacy. For this, well-known techniques are used, in particular, retrospective analysis.

Most often, differential equations are used to model complex systems that describe the dynamics of changes in the states of such systems. As a rule, this is a system of first-order equations having the form [89]

$$dx_i / dt = f_i\left(X, a\right),$$

where $X = \left(X_1, ..., X_n\right)$ is the vector of variables characterizing the state of the social system ; a is the vector of system parameters ; t - time.

The solutions of the reduced system of equations are usually represented as trajectories in the phase space. If we fix the values of all parameters, i.e. choose a point in the parametric space, then the solutions of the reduced system of differential equations will depend only on the initial conditions. However, for a qualitative approach, it is not so much particular solutions that are important, but the most complete description of the behavior of the system in the entire dynamic space [90].

This general picture will predominantly depend on the values to which the solutions aim at $t \to \infty$ or at $t \to -\infty$. The most important asymptotic solutions of this type for the subject area under consideration are stationary points and limit cycles. In this case, of greatest interest is a special type of system stability: stability with respect to changes in system parameters. A system whose general dynamic character does not change with small changes in parameters is called rough or “hard” (eng. - *hard).* Otherwise, the systems are called "soft".

The first step in investigating the above system of differential equations is to determine the stationary points, i.e. solution of the system of equations:

$$f_i\left(X, a\right) = 0 .$$

The second step of research is to determine the nature of singular points. To do this, we pass to new variables - deviations from the coordinates of a stationary point:

$$u_i = x_i - x_i^0 .$$

Near the stationary point, the right-hand sides of the initial equations of the system can be expanded into a Taylor series:

$$\dot{u}_i = \left.\frac{\partial f_i}{\partial x_j}\right|_{x_m^0} u_j + \left.\frac{\partial^2 f_i}{\partial x_k \partial x_l}\right|_{x_m^0} u_k u_l + \ldots$$

Because near x_i^0 performed: $u_i << 1,$ then in many cases we can restrict ourselves to the study of a linear system:

$$\dot{u}_i = \left.\frac{\partial f_i}{\partial x_j}\right|_{x_m^0} u_j = a_{ij} u_j .$$

In qualitative theory, this study is reduced to determining the eigenvalues of the coefficient matrix a_{ij}, i.e. to the solution of the equation

$$\left|a_{ij} - \delta_{ij}\lambda\right| = 0.$$

In the case when all eigenvalues λ_i are different and have non-zero real parts ($\operatorname{Re}\lambda_i$), there is a "rough" stationary point. If all $\operatorname{Re}\lambda_i$ are non-zero, then the following theorems are used:

1. If all $\operatorname{Re}\lambda_i < 0$, then the stationary point is asymptotically stable.

2. If at least one $\operatorname{Re}\lambda_i > 0$, then the stationary point is unstable.

The question of the stability of a stationary point can also be solved on the basis of the Hurwitz criterion, without direct calculation of the eigenvalues. For this, the characteristic equation is rewritten in the form

$$a_0\lambda^n + b_0\lambda^{n-1} + a_1\lambda^{n-2} + b_1\lambda^{n-3} + \ldots = 0,$$

where a_0 =1.

The Hurwitz matrix is a matrix of n-th order:

$$H = \left\| \begin{matrix} b_0 & b_1 & b_2 \ldots b_{n-1} \\ a_0 & a_1 & a_2 \ldots a_{n-1} \\ 0 & b_0 & b_1 \ldots b_{n-2} \\ 0 & a_0 & a_1 \ldots a_{n-2} \\ 0 & 0 & b_0 \ldots b_{n-3} \\ . \quad . & . & . \quad\quad . \end{matrix} \right\|.$$

The minors of the matrix H (from the first to the n-th order) are called the Hurwitz determinants. The Hurwitz criterion is as follows: in order for all $\operatorname{Re}\lambda_i < 0$, a necessary and sufficient condition is the positivity of all Hurwitz determinants. As a result, a general study of the nature of the singular points of the original system can be based on reducing its matrix to the Jordan form.

In the simplest case, for a second-order system, the initial equation can be written as

$$\lambda^2 - \sigma\lambda + \Delta = 0,$$

Where $\sigma = a_{11} + a_{22,}\ \Delta = a_{11}a_{22} - a_{12}a_{21}$.

Then

$$\lambda_{1,2} = \frac{\sigma}{2} \pm \frac{1}{2}\sqrt{\sigma^2 - 4\Delta}.$$

In this simplest example, we confine ourselves to considering only coarse stationary points. If none of the parameters $\sigma,\ \Delta$ is equal to zero, then the qualitative picture of the phase space in the vicinity of the stationary point depends only on the linear terms and the stationary point is coarse. There are three rough equilibrium positions:

1. Stationary point of the "node" type (Fig. 31, *a)*: eigenvalues λ_i are real, of the same sign, $0 < \Delta < \dfrac{\sigma^2}{4},\ \ \sigma \neq 0.$ At $\sigma < 0$ the node is stable, at $\sigma > 0$ - unstable.

2. Stationary point of the "saddle" type (Fig. 31, *b)*: eigenvalues λ_i are real, different signs, $\Delta < 0,\ \ \sigma \neq 0.$

3. Stationary point of the "focus" type (Fig. 31, *c)*: complex eigenvalues $\lambda_i, \Delta > \dfrac{\sigma^2}{4},\ \ \sigma \neq 0.$ At $\sigma < 0$ — stable focus, motion near the stationary point has the character of damped oscillations. At $\sigma > 0$ the focus is unstable, oscillations of increasing amplitude occur.

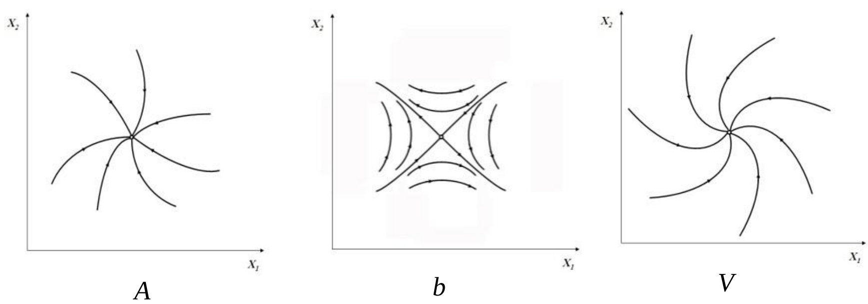


Figure 31. Phase portraits of coarse singular points

An analysis of phase trajectories makes it possible to draw a conclusion about the nature of the evolution of the system, to determine the areas of its deterministic behavior and the areas of bifurcations (i.e., the areas of parameter values at which instability occurs and the form of solutions to the equation describing the behavior of the system changes).

Complex systems often have several stable states (attractors), in one of which they sooner or later find themselves. In these cases, the paths of evolution are not discrete: only a certain set of paths corresponding to attractors is possible (Fig. 32). At the same time, transitions from one attractor to another cannot occur spontaneously, for this it is necessary to change the external conditions or properties of the system. In particular, information operations are used for this in social systems.

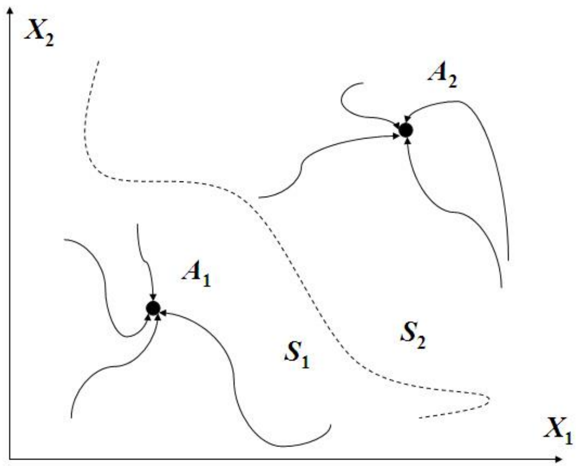


Figure 32. An example of a fragment of the phase space of a dynamic system (a plane cut along two coordinates X_1 and X_2 (A_1 and A_2 are attractors, S_1 and S_2 are the corresponding areas of attraction of attractors)

It is known from synergetics that the transition of a system from one state to another goes through a phase of chaos when the structure-forming processes are weakened. In information systems, chaos provides the system with an initial set of options for further development - attractors. During periods of chaos, crises, when opportunities arise multivariant development, information operations acquire a decisive role. During these periods, social and information systems are most vulnerable to information influences that can play a decisive role in the further development of events [89], the choice of one or another attractor.

Thus, it can be stated that at different stages of evolution, information systems have different sensitivity and vulnerability to external influences - information operations that become most dangerous during periods of chaos - crises. In these cases, even not too intense information impacts can set the direction of the system development and affect the nature of its subsequent evolution. Of course, as well as for the effective conduct of information operations, it is necessary to know the structure of the available attractors, as well as to master a set of techniques for transferring the system from one attractor to another.

Consider specific nonlinear models of information systems dynamics, starting with the simplest growth model ($\dot{x} = kx$), which was proposed by Malthus to calculate the dynamics of the population of the Earth. This model leads, as shown above, to an exponential increase in the population x over time; it can be applied, for example, to the development of information systems at an early stage. The solution of the above equation, as is known, is the exponential.

Obviously, no real process, neither physical nor informational, can develop according to the exponential law for an unlimited time. Indeed, starting from a certain moment, the dependence too quickly tends to infinity, which, for obvious reasons, is not realized in nature. Therefore, we have to admit that sooner or later, and sooner rather than later, some catastrophe will occur, which will change the nature of the dependence and return it within the limits of the permissible range of values.

In more or less stable systems, there is always an element of self-consistency, due to which, over significant time intervals, the dependence of the dynamics of their development is described by more complex equations containing feedback. Therefore, the nature of the dependence changes over time, and the changes occur in a non-trivial way. This results in the following typical cases:

- dependence reaches saturation , and the system goes into a static (or, possibly, homeostatic) state;

- the dependence has a local maximum, followed by a decrease (including to zero);

- an oscillatory mode is established (usually damped, but self-oscillating is also possible).

In real life, as a rule, dynamic systems have fairly effective feedbacks that allow you to correct the nature of the processes occurring in them and thereby keep them within certain limits. Information operations, correcting these feedbacks at certain periods of the evolutionary process, can quite effectively influence the behavior of the entire system.

As in the case of information flows, a widely used generalization of Malthus's law, known as the logistic model, is used.

5.4. Interaction of information systems

The logistic equation describes the dynamics of one information system interacting only with the surrounding information space. In the theory of population dynamics, a classification of various forms of such interaction between populations has been developed [91–93] , in our case, information systems.

The main information systems include the following:

- neutralism (lack of direct influence of populations on each other);
- competition (mutual suppression of populations);
- ammensalism (unilateral suppression of one population);
- predation (destruction by individuals of one population of individuals of another);
- symbiosis (productive coexistence of populations).

In the dynamics of interacting populations, two categories of influences are distinguished, which differ in their temporal nature:

- phase (single);
- parametric (constant).

The logistic model makes it possible to quite satisfactorily describe the dynamics M interacting populations. In the general case, this is done using the system of equations already given for the case of information flows, but with a slightly different meaning of the parameters:

$$\frac{dn_i(t)}{dt} = n_i(t)[p_i - \sum_{j=1}^{M} q_{ij} n_j(t)],$$

$$n_i(0) = n_{0i}.$$

The type of process described by this system of equations is determined by the value and sign of the coefficients p_i and q_{ij} . It should also be taken into account that in each equation, the diagonal terms $n_i(t)n_j(t)$ describe intraspecific interactions, while the cross terms describe $n_i(t)n_j(t)$ interspecific ones.

In other words, the diagonal terms describe the impact of the external environment on the population, including the depletion of available resources, and the cross terms describe the impact of one population on another (positive values correspond to a favorable effect, negative values correspond to an unfavorable one). The coefficients p_i have the meaning of the growth rates of the respective populations in the absence of interaction.

An important point is also the behavior of the population for given values of the parameters and in the absence of interaction (for example, its growth is limited by itself).

The above system of equations, in principle, can describe a wide range of dependencies, and this, in a certain sense, is a problem, since, if desired, anything can be "pulled out" from its solutions. Therefore, working with it requires a balanced and responsible attitude.

However, decisions that characterize real processes usually refer to one of the following modes:

- stationary;
- self-oscillating;
— quasi-stochastic.

As a rule, these regimes manifest themselves to the full extent on sufficiently large (not necessarily infinite) time intervals. But the transient processes that precede the establishment of a certain regime are exclusively polymorphic, their behavior can largely determine the subsequent dynamics. It is these processes that can be considered as the main object of information operations in the case of planning social procedures.

Below are the results of modeling the frequency characteristics of some information processes within the logistic model. The undoubted advantage of this model is that it combines the simplicity of the initial formulations with the flexibility in setting tasks.

The above description of population dynamics within the framework of the logistic model was first formulated for biological systems, but has now been extended to other areas of research, including information processes.

The given system of logistic equations allows one to describe the dynamics of any number of populations, both interacting with each other and those in isolation. But to understand its basic patterns, it is enough to limit ourselves to a small number of them. Even the study of the behavior of two interacting populations makes it possible to trace the general patterns of their dynamics, at least qualitative ones. Practice shows that this is quite enough for reasonable forecasts.

Below we study the dynamics of three information systems, of which two are considered basic in the sense that their behavior should illustrate the points of interest to us, and the third (additional) is introduced to demonstrate the role that the general information context can play in this case. At the same time, three main options for the interaction of information systems with systems are studied: competition, predation, and symbiosis.

Since the analytical solutions of the above system of equations, in cases where they can be constructed, turn out to be cumbersome and difficult to analyze, numerical methods are initially used, especially since the graphical form of presenting the results in this case is the most convenient and illustrative. Since we are interested in the qualitative behavior of dependencies, the results below will be presented in conventional units.

5.4.1. Dynamics of the "Competition" type

Competition is a form of interaction between populations, in which they mutually suppress each other (in the simplified model under consideration, due to the limited common resource base). The main feature of competition is that competing populations do not directly affect each other. Interaction is carried out indirectly - by ousting each other from the area of limited resources. In this case, complete suppression of one of the populations is possible, as a result of which it disappears. It is competitive relations that pose a real danger to information systems, and it is they that are most characteristic of the main participants in information processes that are in information confrontation.

Depending on the conditions in which the interacting information systems are located, and the values of the parameters determining the dynamics, both various equilibrium states of the system and the mechanisms for achieving them are possible.

Studies of the stability of systems, populations in particular, to external influences are carried out on the basis of competition modeling, to which a large number of works are devoted. In particular, in [94], a system of non-linear differential equations is presented that describes the change in the balance of forces of the opposing sides as a result of competitive struggle.

In this model, it was shown that the nature of the evolution and self-organization of systems depends to a decisive extent on the following circumstances (conditions):

— stationarity of the system functioning;

— closedness or openness of the system;

— resource provision of the system.

Below are the most characteristic, in our opinion, cases of the dynamics of information systems in competitive relations. Competition corresponds to the above system of logistic equations with positive values of the coefficients q_{ij} .

Equilibrium coexistence without external influences

We will assume that the interaction between the two main forces, on the one hand, and the third force, on the other, is mainly reduced to the mutual limitation of the resource base. The main rivalry takes place between the main forces.

With sufficiently small values of the coefficients q_{ij} describing the influence of one competing information system on another; at sufficiently large values of their growth rates , each system reaches an equilibrium state and stabilizes in it. Depending on the values of other parameters and the initial size, the information system in the process of reaching this state can either increase or decrease.

As we see from fig. 33, both main information systems reach an equilibrium value, but at the same time, the graph of one of them increases, while the other decreases.

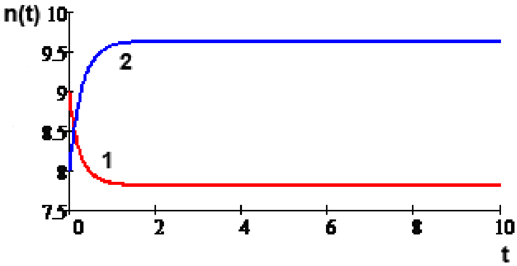


Figure 33. Equilibrium coexistence of information systems 1 and 2 in their own niches

It may seem that the above results are obvious, but if we consider that we are talking about competitive relations, this is by no means the case. It turns out that both competing information systems may find themselves in an equilibrium state, from which they themselves will not be able to get out under any circumstances.

Complete suppression of one information system by another

With a significant increase, q_{ij} the number of one of the information systems is reduced to zero (Fig. 34). If the values of the coefficients q_{12} and q_{21} are close, then the situation becomes unstable in the sense that whether the information system is suppressed or not depends on small deviations in the values of other parameters.

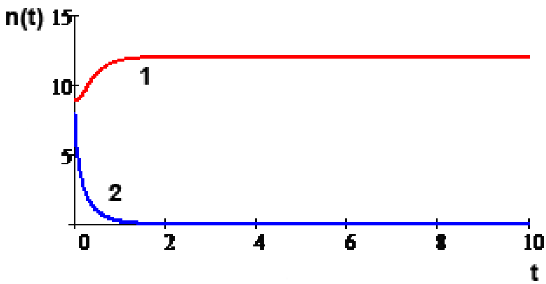


Figure 34. Complete suppression of one information system 2 by information system 1 (here t is *time, n (t) is the size of the information system)*

The suppression of one of the competing information systems by another is a scenario that is most often perceived as natural and logical.

Therefore, the results presented as such are not of particular interest. Another thing is interesting : under the chosen conditions, the winning information system has not exhausted the "freed up resources". Its increase is insignificant compared to the losses of the losing information system.

Thus, victory in competition does not mean automatic support from the information resources of the competing party.

Equilibrium coexistence through a third force

In previous cases, the direct influence of the third force on the main ones was neglected. However, if the third force in one way or another "feeds" one of the main forces, i.e. is an information operation of the first information system relative to the second, then their dynamics can change dramatically (Fig. 35). For example, a coexistence scenario is possible.

The equilibrium coexistence of competing main information systems, which arises due to the positive impact on one of them of an additional information system, reflects such an extremely important situation as the impact on interacting information systems of a new context, which can be both positive and negative. In both cases, such an influence can effectively compensate for strong and weaknesses of the main competitors, leading to seemingly unexpected scenarios.

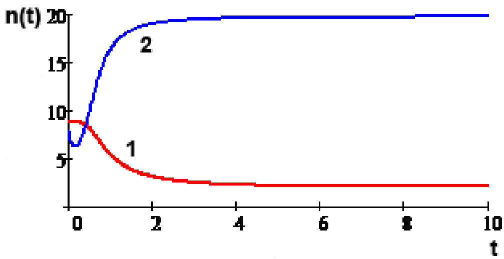


Figure 35. Equilibrium coexistence of information systems 1 and 2 by a third force

One of the key issues is the stability of solutions with respect to the numerical values of the parameters. In a number of cases, the behavior, including the qualitative one, of the solutions of the above system of equations depends very strongly on them. Let us illustrate this by the example of the dependence of solutions on the growth rate (Fig. 36).

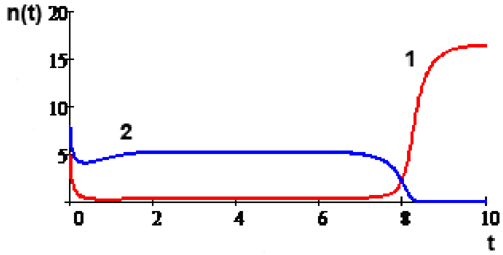
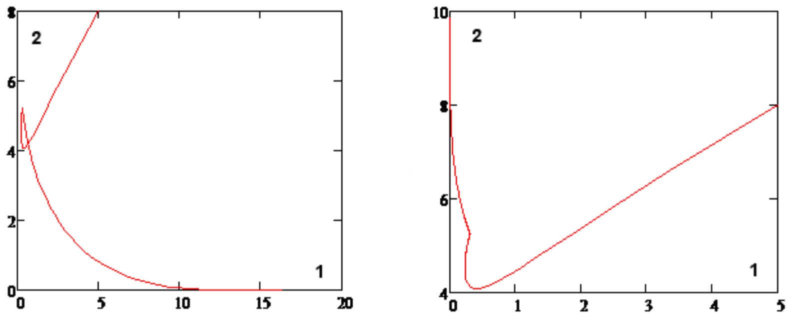


Figure 36. Dependence of the size of information systems 1 and 2 on their growth rate

The graphs presented differ in minor changes in the coefficient p_2. As we can see, in this case the behavior of the reduced curves differ not only quantitatively, but also qualitatively. This is more clearly shown in the phase portraits (Fig. 37), corresponding to information systems, the behavior of which is reflected in Fig. 3. 36.

Thus, the dynamics of information systems can significantly depend on small changes in their growth rate. This means that an information system can outperform the competition due to a small advantage in this feature.



A *b*

Figure 37. Phase portraits of size dependencies information systems 1 (a) and 2 (b) on the rate of their growth

We note the non-trivial behavior of the first dependence: at first, the second information system dominates, and the first one is suppressed, but then they change places. Thus, when two information systems compete in the presence of a third, situations are possible when one of them, which at first was the undisputed leader, is eventually forced out to the periphery.

5.4.2. Dynamics of "Predation" type

Absorption dynamics

Predation is often viewed as a form of competition in which the components of one information system directly absorb the components of another. In contrast to the case of ordinary competition, here the relationship of information systems is asymmetric: one plays the role of a predator, the other is a victim, and they cannot be swapped.

Predator-prey relations by themselves never lead to the suppression of one information system by another: predators cannot reproduce their population without prey - their food (as literally and figuratively). The reduction in the population of prey, in turn, causes a reduction in the population of predators, since some of them are left without means of subsistence. But the reduction in the number of predators leads to a decrease in the external impact on the prey, and they begin to recover. Therefore, this case is characterized by various oscillatory regimes.

Naturally, the influence of a third force can significantly change the typical picture. This type of interaction of information systems is the most complex and, at the same time, the most interesting. The main reason is that the mechanisms of the second group play the main role here.

Indeed, in the case of interaction between biological species, a decrease in the number of prey, starting from a certain moment, causes a decrease in the number of predators, since this reduces their resource base (available volumes of food). In information systems, at the level of their sizes, nothing similar happens. On the contrary, the components of one information system, "eaten" by another, mechanically increase the number of predators, and without any restrictions. An information system that has absorbed the rest becomes an information monopoly for a certain time. Therefore, the mechanisms of the first group by themselves in this case can only lead to trivial effects of mechanical expansion of one force at the expense of others.

With the mechanisms of the second group, the situation is much more complicated. Often the weight of one information system is based on the exploitation of the resources of another. For example, the propaganda of competing forces can be built on opposites that exclude each other, but at the same time give meaning to each other. In such cases, to maintain activity, an adversary is needed, with whom it would make sense to argue. The real "information weight" is determined by superiority in this controversy. Another example is the use of the "image of the enemy".

It is also possible to systematically exploit one information system of meaningful developments created by another.

Under such scenarios, it is quite possible to speak of a predator - prey relationship between information systems. Let us analyze several situations characteristic of predation relations. Predation corresponds to the above system of equations with negative values of the coefficients p_i and q_{ij} for the predator and positive values for the prey.

Impact of the information operation

If the third force (information system) has a weak effect on the main information systems (does not participate in the consumption of the corresponding resources), then we have the usual picture of predation (Fig. 38).

There are typical fluctuations in the size of both information systems, and fading. In sufficiently large time intervals, both dependences tend to a state of some equilibrium, but the oscillation amplitude never becomes equal to zero.

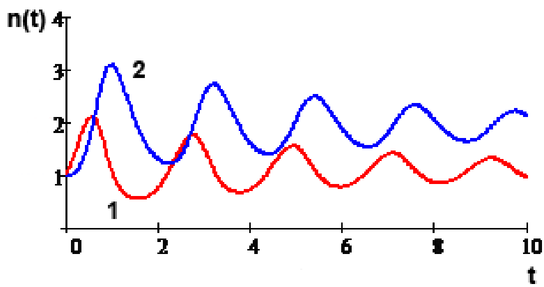


Figure 38. Weak influence of the third force for information systems 1 and 2

Depending on the values of the parameters, the curves can have different phase ratios, which sometimes leads to interesting effects, which are the subject of a separate study.

For reasons of clarity, let us assume that the initial numbers of both basic information systems are equal. The oscillatory regime characteristic of the victim-relationship in the behavior of the dynamics of information systems has indeed been observed in a number of analytical studies.

The next two relationships show us how the dynamics of predator and prey can be influenced by a third force (Fig. 39). The given solutions differ from each other by the value of the coefficient q_{13}. It is equal to -1 in the first case (positive impact on the first of the main information systems) and 1 in the second (negative impact on it).

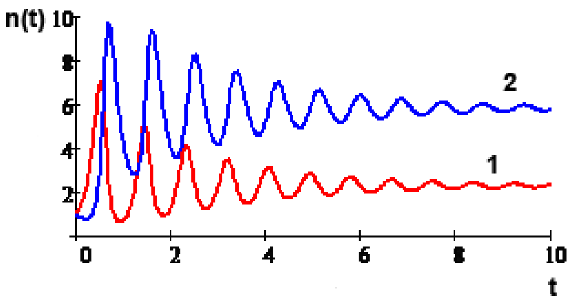


Figure 39. The influence of a third force on information systems 1 and 2

With a positive impact of the third force, significant changes in the nature of oscillatory processes are also observed, in particular, their frequency increases, but at the same time, the decay time is reduced (external stabilizing effect). It is also important that information systems can repeatedly change places in size. From the point of view of an outside observer, such a process may look strange and disturbing, but it quickly normalizes and passes into a (quasi) equilibrium state.

5.4.3. Dynamics of the "Symbiosis" type

The symbiosis of information systems occurs when, for some reason, they either do not interfere with each other, or support each other. In this case, most often there is no need for additional information operations, and the intervention of a third force turns out to be ineffective.

In some cases, populations (information systems in symbiotic relationships) have a positive effect on each other, helping to survive in a tough struggle with other systems. However, one should not think that symbiosis always represents a peaceful and blissful coexistence. In reality, relationships can be rigid and even antagonistic. It's just that such information systems may not have a real opportunity to influence each other through information mechanisms.

Symbiosis is perhaps the least interesting both in theoretical and applied terms. Indeed, the dynamics of information systems that are in a symbiotic relationship is not much different from the dynamics of non-interacting systems. However, the very fact that interacting systems can be in such a state is important. However, for the sake of completeness, it is included in the overall picture.

Symbiosis assumes that for all information systems participating in it, the coefficients p_i are positive, and q_{ij} negative.

We also note that information systems between which there are symbiosis relations quickly and almost simultaneously reach their equilibrium states. Thus, this type of system interaction can be called the most static.

5.5. Time series analysis

The basis for successful modeling of information processes and forecasting their results is taking into account the relationship of events with the information space, in particular , with its most dynamic and modern part - the information resources of the web space. The task of studying the properties of the information space and information systems is multifaceted, it involves the active use of methods for analyzing complex dependencies, time series, which allow a deeper understanding of the specifics of a particular subject area.

Let us consider the possibilities of modern analytical tools on the example of the study of thematic information flows of web publications collected from the Internet by the InfoStream system [95].

As a request to the InfoStream web resource monitoring system, expressing the topics of the studied information array related to the development of crisis phenomena in Ukraine in 2008 [96], the following expression was chosen:

(Parliament ~ Riz)|(Political~Crisis)|(Finance~Crisis)|(Economic~Crisis)

Information flows coming from more than a thousand Ukrainian network information resources were studied, among which the leaders in the number of publications relevant to the request were such authoritative sources as Ukrinform, UNIAN, RBC-Ukraine , Radio Liberty, Korrespondent.net, Glavred, etc. The retrospective period of the study was the entire year of 2008, i.e. 366 days. During this period, the InfoStream system covered over 12 million network documents. As a result of a search on request, which took into account all the main aspects of the crisis phenomena, 57245 relevant documents were found. Based on the processing of these data, quite complete pictures of experimental data were obtained - time series for a given period. On fig. 40 shows a graph of changes in the number of relevant thematic publications by day in 2008.

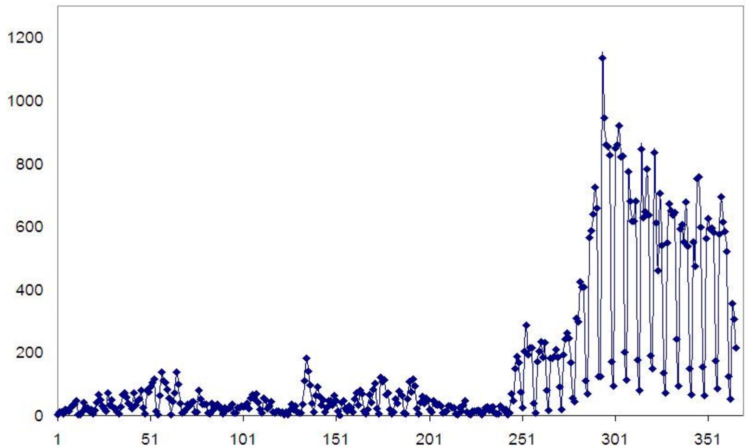


Figure 40. Dynamics of the number of thematic publications by days of 2008 (total 57245 publications)

The above graph allows you to see weekly fluctuations in the volume of publications (on weekends, for example, significantly fewer documents are published online than on weekdays). It can also be seen from the graph that on approximately the 250th day of the year, the total number of messages on crisis issues increased sharply (the parliamentary crisis intensified).

5.5.1. Correlation analysis

One of the main methods of modern analysis of measurement series is correlation analysis. Let us dwell in more detail on the formalism of correlation analysis.

If we designate X_t member of a series of measurements (for example, the number of emails received per day t, $t = 1, \ \ldots, \ N$), then the autocorrelation function for this series X with a "measurement window" in k days is defined as follows:

$$F(k) = \frac{1}{N-k} \sum_{t=1}^{N-k} \left(X_{k+t} - m \right) \left(X_t - m \right),$$

where m is the mean value of the series X . Autocorrelation coefficients for series of measurements X long N with a measurement window equal to k , are calculated by the formula

$$R(k) = \frac{F(k)}{\sigma^2},$$

where is $F(k) -$ the autocorrelation function; σ^2 - dispersion.

An important property of the autocorrelation function is the ability to identify harmonic components, as well as the self-similarity of the original process. It is known that this function has the property that if there is a hidden periodic component in a series of measurements, then its value tends asymptotically to the square of the mean value of this series. In addition, the autocorrelation function of the periodic series is also periodic, contains the fundamental frequency and harmonics. If a series of measurements X is the sum of some meaningful component N and a sinusoidal signal S , then the autocorrelation function of the series X includes an explicitly expressed periodic component [97].

The graphical representation of the autocorrelation coefficient for a number of observations corresponding to the dynamics of the thematic information flow of web publications considered above indicates the invariance of the correlation properties by day of the week (Fig. 41), and the trend indicates the possible self-similarity of the original time series.

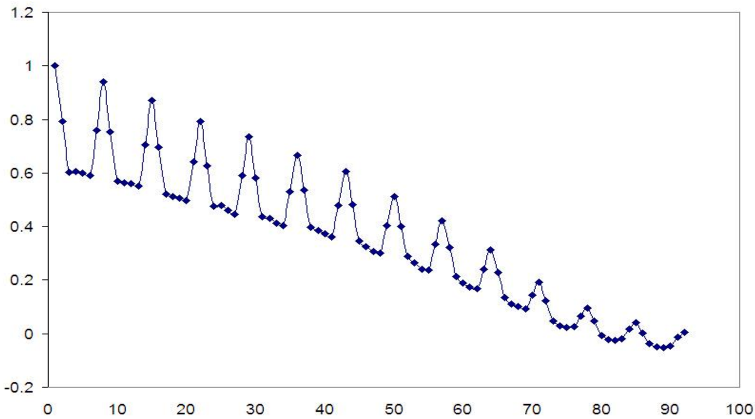


Figure 41. Autocorrelation coefficients of a series of observations R (k) (y-axis) depending on k (abscissa)

5.5.2. Wavelet analysis

The basis of wavelet analysis [98, 99] is the wavelet transform, which is a special type of linear transformation whose basis functions (wavelets) have specific properties. A wavelet (small wave) is a function that is concentrated in a small neighborhood of a certain point and decreases sharply to zero as it moves away from it both in the time and frequency domains. There are various wavelets with different properties. At the same time, all wavelets have the form of short wave packets with zero integral value, localized on the time axis, which are invariant to shift and scaling.

Two operations can be applied to any wavelet:

- shift, i.e. moving the area of its localization in time;
- scaling (stretching or shrinking).

The main idea of the wavelet transform is that the non-stationary time series is divided into separate intervals (the so-called observation windows), and on each of them the scalar product (a value that characterizes the degree of closeness of two patterns) of the studied data is calculated with different shifts of some wavelet on different scales. The wavelet transform generates a set of coefficients that represent the original series. They are functions of two variables: time and frequency, so they form a surface in three-dimensional space. These coefficients show how the behavior of the process at a given point is similar to a wavelet on a given scale. The closer the type of the analyzed dependence in the vicinity of a given point to the wavelet type, the greater the absolute value of the corresponding coefficient. Negative coefficients show that the dependence is similar to the "mirror reflection" of the wavelet. The use of these operations, taking into account the locality property of the wavelet in the time-frequency domain, makes it possible to analyze data on different scales and accurately determine the places of their characteristic features in time.

Wavelet technology makes it possible to detect single and irregular "bursts", sharp changes in the values of k quantitative indicators in different periods of time, in particular, the volume of thematic publications in the web space. In this case, moments of the occurrence of cycles, as well as moments when periods of regular dynamics are followed by chaotic oscillations, can be detected.

The considered time series can be approximated by a curve, which, in turn, can be represented as a sum of harmonic oscillations of different frequencies and amplitudes. In this case, oscillations that have a low frequency are responsible for slow, smooth, large-scale changes in the values of the original series, and high-frequency ones are responsible for short, small-scale changes. The stronger the value described by this regularity changes at a given scale, the greater the amplitude of the component of the corresponding frequency. Thus, the time series under study can be considered in the time-frequency domain, i.e. in the field of research, patterns that describe the process depending on both time and frequency.

A continuous wavelet transform for a function $f(t)$ is built using continuous scaling and translations of the selected wavelet $\psi(t)$ with arbitrary values of the scale factor a and the shift parameter b :

$$W(a,b) = \left(f(t), \psi(t) \right) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \psi^* \left(\frac{t-b}{a} \right) dt .$$

The obtained coefficients are presented in graphical form as a map of the transformation coefficients, or a scalogram. On the scalogram, the wavelet shift (time axis) is plotted along one axis, and the scales (scale axis) are plotted along the other axis, after which the points of the resulting scheme are colored depending on the values of the corresponding coefficients (the larger the coefficient, the brighter the colors of the image). On the scalogram, all the characteristic features of the original series are visible: the scale and intensity of periodic changes, the direction and value of trends, the presence, location and duration of local features.

For example, it is known that the combination of several different fluctuations can have such a complex shape that it does not allow the analyst to identify them. The periodic changes that occur for the values of the wavelet transform coefficients on a certain continuous set of frequencies look like a chain of "hills" with vertices located at the points (along the time axis) where these changes reach the greatest values.

Another important indicator is a pronounced trend in the dynamics of the time series (trend), regardless of periodic fluctuations. The presence of a trend may not be obvious by simply looking at the time series, for example, if the trend is combined with periodic fluctuations. The trend is reflected on the scalogram as a smooth change in brightness along the time axis simultaneously on all scales. If the trend is increasing, then the brightness will increase, if it is decreasing, it will decrease.

Another important factor to consider when analyzing time series is local features, i.e. possible sharp, spasmodic changes in the characteristics of the original series. Local features presented on the wavelet transform scalogram as lines of a sharp brightness drop that emanate from a point correspond to the time of the jump. Local features can be both random and systematic, while "masking" periodic dependencies or a short-term trend. An analysis of local features makes it possible to recover information about the dynamics of the initial process and, in some cases, to predict such situations.

Thus, each of the main dynamics factors has its own characteristic reflection on the scalogram, and all analytical information is presented in a visual and convenient form for studying. Due to the visual presentation of the results in the form of a scalogram, sometimes one glance is enough to see the most significant factors on it [100, 101].

On fig. 42 shows a scalogram - the result of a continuous wavelet analysis (Gaussian wavelet) of the time series corresponding to the process considered above.

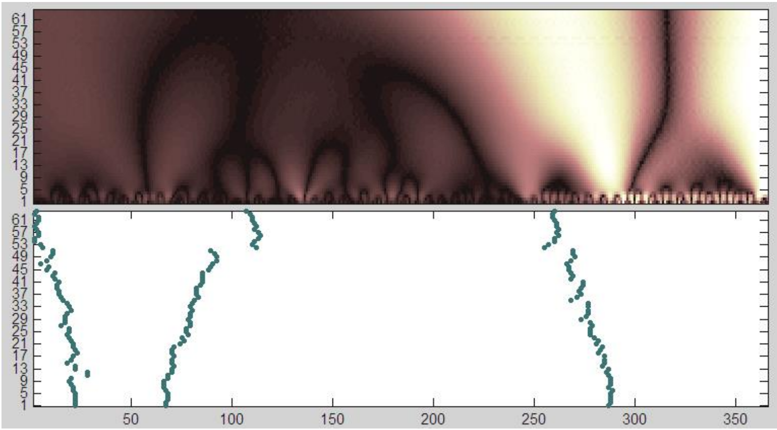


Figure 42. The result of wavelet analysis (continuous wavelet transform): top - wavelet scalegram;
below - lines of local maxima (skeleton)

The above example shows that wavelet analysis can detect not only obvious anomalies in the series under study, but also critical values that are hidden behind relatively small absolute values of the elements of the series. For example, on the skeleton, the highest values were noted not only on the 250th day, but implicit extrema were also shown (on the 25th and 75th days).

Certainly, financial and economic factors have a direct impact on social processes. On fig. 43 shows the dynamics of changes in the exchange rate of selling US dollar cash in Ukrainian banks during 2008. On fig. 44 shows an example of applying wavelet analysis to this series. The shift of the last line of local trends on this scalogram in comparison with the previous one (see Fig. 42) indicates that the cash rate is a derivative of general crisis phenomena.

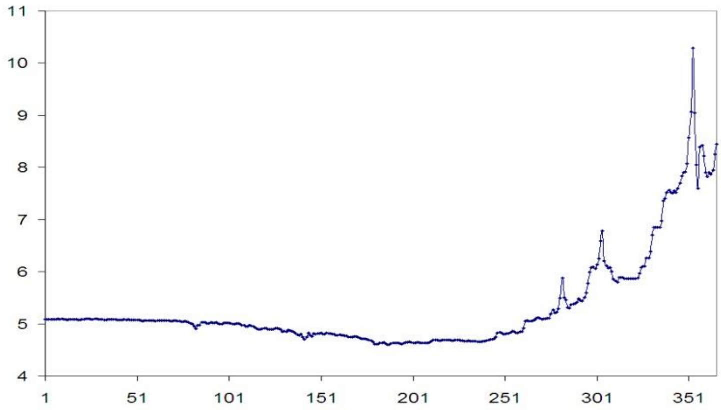


Figure 43. Dynamics of changes in the exchange rate of the cash US dollar in hryvnia during 2008 г.

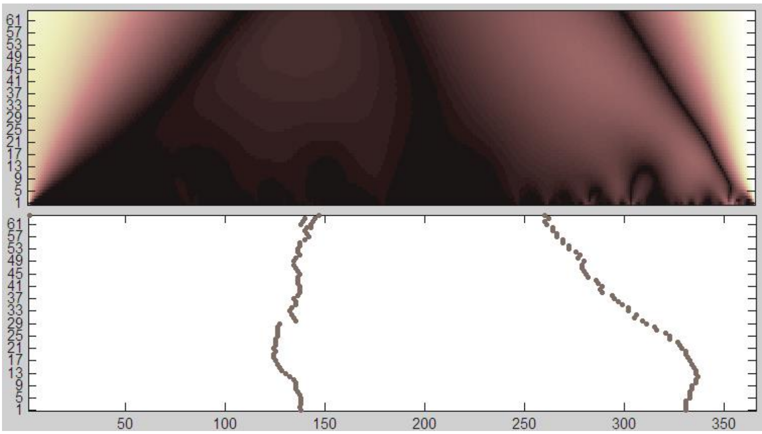


Figure 44. The result of the wavelet analysis of the series corresponding to the course cash US dollar: above - wavelet scalegram; below - lines of local maxima (skeleton)

5.5.3. Fractal Analysis: R/S Analysis

The theory of fractals [102] is widely used as an approach to the study of observational series, which makes it possible to obtain important characteristics of the corresponding processes without going into a detailed analysis of their internal structure.

Self-similarity analysis of time series can be considered as a technology designed to carry out analytical studies with forecasting elements, suitable for extrapolating the obtained dependencies.

The most important characteristic of series with chaotic behavior is the fractal dimension, which in many cases can be calculated using the so-called R/S analysis. More precisely, it is not the fractal dimension itself that is calculated, but the Hurst exponent, which is related to it by a simple relation. R/S - the analysis is based on the analysis of the normalized scatter - the ratio of the scatter of the values of the series under study R to the standard deviation S [103].

To study the fractal characteristics of time series $F(n)$, $n=1,...,N$, composed of the number of messages belonging to them , we studied the value R/S, where R is the so-called range:

$$R(N) = \max_{1 \le n \le N} X(n,N) - \min_{1 \le n \le N} X(n,N),$$

a S is the standard deviation:

$$S = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(F(n) - \langle F \rangle_N\right)^2};$$

$$\langle F \rangle_N = \frac{1}{N}\sum_{n=1}^{N} F(n);$$

$$X(n,N) = \sum_{i=1}^{n}\left(F(i) - \langle F \rangle_N\right).$$

For a fairly wide class of series, the dependence R/S has an exponential trend, i.e. we can talk about the relationship

$$R/S = \left(\frac{N}{2}\right)^H,$$

Where H is the Hurst exponent, which, under certain additional conditions, is associated with the Hausdorff (fractal) dimension D wait with the formula: $D+H=2$.

The main condition under which the Hurst exponent is associated with the fractal dimension in accordance with the above formula was determined by E. Feder: “... cells are considered, the sizes of which are small compared to both the duration of the process and the range of function change; therefore, the relationship is valid when the structure of the curve describing the fractal function is examined with high resolution, i.e. in the local limit. Another important condition is the self-affinity of the function. Without going into details, we note that, for example, for information flows, this property is interpreted as self-similarity resulting from the processes of their formation. At the same time, not all information flows have these properties, but only those that are characterized by sufficient power and iteration during formation.

On fig. 45 presented values R/S for a number of publications by day in 2008, which corresponds to the above query. Obviously, the nature of the normalized range changes dramatically around the 250th day of the year approximately at the time when the first serious statements were made at the highest level about the financial and economic crisis, i.e. we actually have two different rows - from 1 to 250 and from 251 to 366. As you can see, the normalized range curve for the second row (Fig. 46) is satisfactorily approximated by a straight line on a double logarithmic scale. The slope of this straight line corresponds to the Hurst exponent.

Numerical values H characterize different types of correlation dynamics (persistence). At $H = 0,5$ there is an uncorrelated behavior of the values of the series, and the values $0,5 < H < 1$ correspond to the level of autocorrelation of the series. As you can see, the Hurst exponent for the studied information flow in Fig. 46 corresponds to a value of ~0.89, which confirms the assumption about the self-similarity and iterative processes in the information space. Publications, citations, hypertext and contextual links, etc. generate self-similarity, the presence of a high level of statistical correlation in information flows over long time intervals.

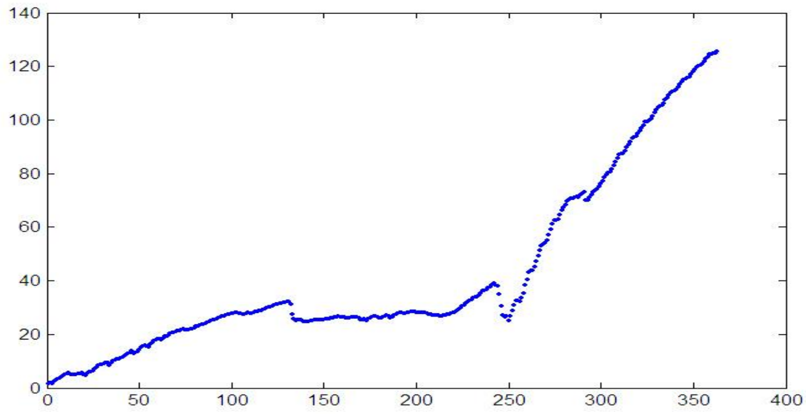


Figure 45. Normalized scatter index (y-axis) for all observation period of the series under study (abscissa)

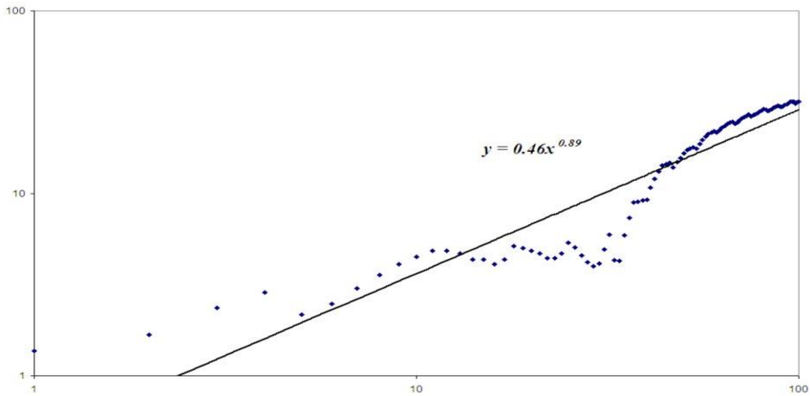


Figure 46. Normalized scatter exponent on a logarithmic scale for the last 120 days of the year

5.5.4. Deviation from a linear trend

The DFA (Detrended fluctuation analysus) method is also most often used to identify the statistical self-similarity of signals [104].

This method is a variant of the analysis of variance of one-dimensional random walks and allows you to explore the effects of long-term correlations in the series that are being studied. Within the framework of the DFA algorithm, the root-mean-square error of the linear approximation is analyzed depending on the size of the approximation region (observation window). Let there be a number of measurements x_t, $t \in 1,...,N$. Let us denote the average value of this series of measurements: $\langle x \rangle = \frac{1}{N}\sum_{k=1}^{N} x_k$. An accumulation series is constructed from the initial series:

$$X_t = \sum_{k=1}^{t}\left(x_k - \langle x \rangle\right).$$

Then the series X_t is divided into time windows of length , a linear approximation $L\,(L_{j,L})$ is constructed using the values $X_{k,j,L}$ of c $X_{j,L}$ inside each window (in turn, $X_{j,L}$ a subset X_t of , $j = 1,..., J$, $J = N / L$ is the number of observation windows), and the deviation of the points of the accumulation series from the linear approximation is calculated:

$$E(j,L) = \sqrt{\frac{1}{L}\sum_{k=1}^{L}\left(X_{k,j,L} - L_{k,j,L}\right)^2} = \sqrt{\frac{1}{L}\sum_{k=1}^{L}|\,\Delta_{k,j,L}\,|^2},$$

where $L_{k,j,L}$ is the value of the local linear approximation at the point $t = (j-1)L + k$.

Here $|\,\Delta_{k,j,L}\,|$, is the absolute deviation of the element $X_{k,j,L}$ from the local linear approximation.

Next, the average value is calculated

$$F(L) = \frac{1}{J}\sum_{j=1}^{J} E(j,L),$$

after which, in the case $F(L) \propto L^{\alpha}$ where α is a certain constant, conclusions are drawn about the presence of statistical self-similarity and the nature of the behavior of the series of measurements under study.

This method was applied to a series of publication count values obtained from the above query. On fig. 47 shows the dependence of the rms approximation error on the length of the approximation segments in a double logarithmic scale.

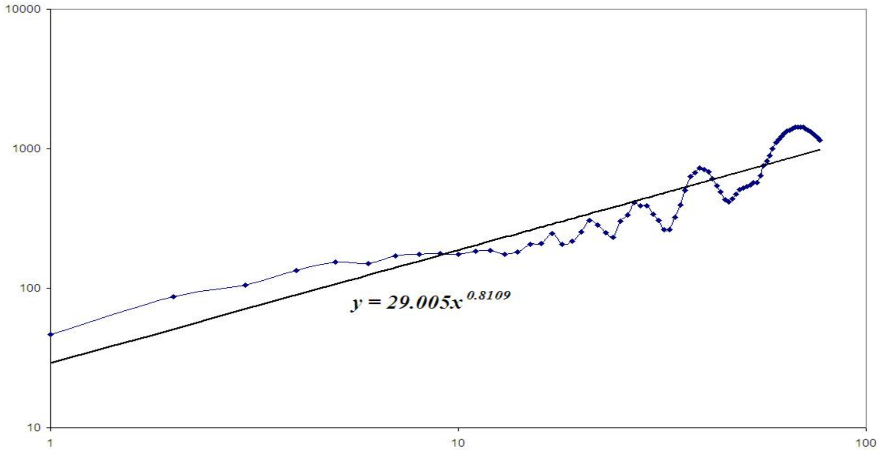


Figure 47. The dependence of the mean square error linear approximation D on the observation window length k

Affinity Dependency $D(k)$ to linear once again confirms the presence of local scaling in the second half of 2008.

5.5.5. Visualization Based ΔL - Analysis

In order to visualize and analyze time series associated with publications in the information space of the Internet, a new method of analysis of variance has been developed to analyze and visualize the state of time series of the intensity of publications on a specific topic [105].

The tasks of identifying and visualizing trends, identifying harmonic components, trends, local features of time series, noise filtering are today solved by methods of fractal, wavelet and Fourier analysis.

As in the DFA method, consider the behavior of the deviation of the accumulation series points from the linear approximation (but in this case the absolute value) $|\Delta_{k,j,L}|$. The construction of the corresponding diagrams of values $|\Delta_{k,j,L}|$, which actually depend on two parameters - L and $t=(j-1)L+k$, is called ΔL - a visualization method. Such a visualization in the form of a "relief" diagram is of particular interest for studying the features of processes that correspond to the original series of measurements.

ΔL-method is quite effective for identifying the harmonic components of the series under study. On fig. 48 shows ΔL - a diagram of a series that corresponds to a sinusoid ($y(i) = \sin(i\pi/7),\ i=1,...,366$). The application of ΔL the -method to a series composed of the number of publications collected by the InfoStream system from the Internet without taking into account thematic division has a pronounced harmonic component (the total number of publications depends on the day of the week), which can be seen in Fig. 49. In addition, this chart shows deviations from the overall trend in publications over the holidays.

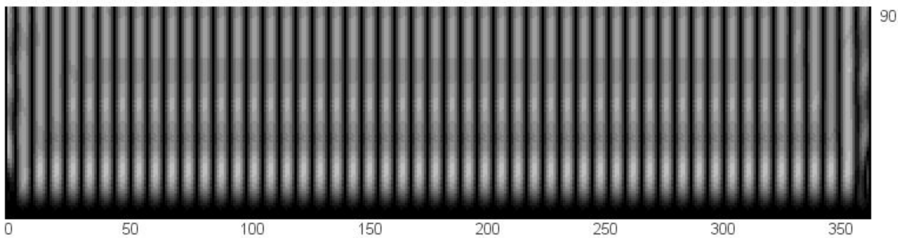


Figure 48. ΔL -sine wave diagram

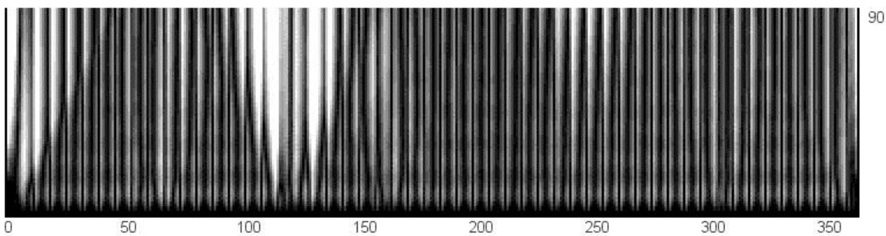


Figure 49. ΔL -diagram of a series of the number of publications collected daily by InfoStream system in 2008

The "relief diagrams" obtained as a result of the ΔL - method (lighter tones correspond to large values of $|\Delta_{k,j,L}|$) resemble scalograms obtained as a result of continuous wavelet transforms. It should be noted that the dark bands in the center of many areas of light shading indicate the "stabilization" of large values of the considered series at a high level.

ΔL - the method is applied to real time series, for example, those that reflect the intensity of publications of this topic on the Internet. On fig. 50 is ΔL a diagram for the time series considered above from the number of messages published every other day on a selected topic on the Internet throughout the year.

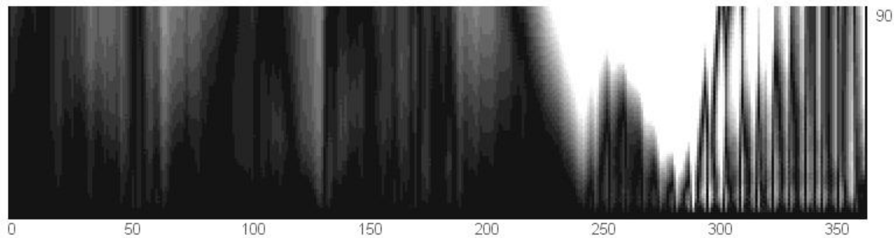


Figure 50. ΔL -time series diagram of intensity thematic publications (abscissa axis - days of the year, the y-axis is the size of the measurement window)

On fig. 51 shows ΔL a diagram of the cash exchange rate of the dollar in hryvnia during 2008. Even more clearly than in the case of wavelet analysis, one can make sure that the most significant deviations in the diagram in this case occur with a certain time delay compared to the diagram of publications on crisis topics.

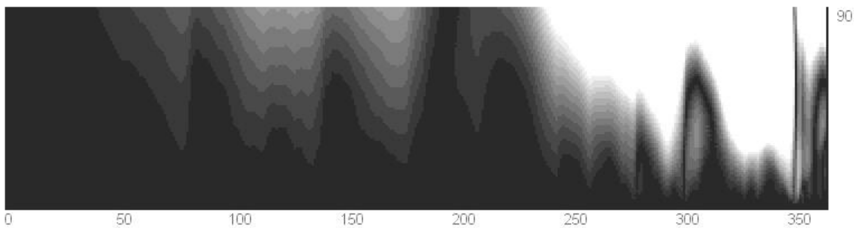


Figure 51. ΔL -time series diagram of cash rates US dollar in hryvnia (abscissa axis - days of the year, the y-axis is the size of the measurement window)

The proposed method for visualizing absolute deviations ΔL - diagrams, as well as the method of wavelet transforms, allows (and, as shown in the example, no worse) to detect single and irregular "bursts", sharp changes in the values of quantitative indicators in different periods of time.

It should be noted that the wavelet transform method can be applied using a variety of wavelets. In the case of applying ΔL the - method, it is not necessary to solve the difficult task of choosing and justifying the application of the corresponding wavelet; unlike the methods of fractal analysis, the proposed approach does not require significant volumes of measurement points. This method is quite simple in software implementation and is based on such a powerful theoretical basis as DFA, and is quite effective in the analysis of time series in such areas as economics and sociology.

5.5.6. Multifractal analysis

The most general description of the nature of self-similar objects is given by the theory of multifractals, which makes it possible to embrace an infinite hierarchy of dimensions and to distinguish homogeneous objects from inhomogeneous ones. The concept of multifractal formalism [106–108] is an effective tool for studying and quantitatively describing a wide variety of complex systems.

In accordance with this formalism, the carrier of a multifractal measure is a set L — the union of fractal subsets L_α, i.e. multifractal can be understood as some kind of union of different homogeneous fractal subsets L_α initial set L, each of which has its own value of the fractal dimension.

To characterize a multifractal set, the so-called multifractal spectrum function $f(\alpha)$ (multifractal singularity spectrum) is used, to which the term “fractal dimension” is fully applicable. The value $f(\alpha)$ is equal to the Hausdorff dimension of a homogeneous fractal subset L_α of the original set L, which makes a dominant contribution to some partition function.

Shown in fig. 52 dependencies refer to the analysis of a numerical series of message intensities that reflect the problems of using anti-virus software (the dynamics of publications in Internet news of messages on a given topic), as well as a series received on a refined topic (the original request was expanded with the word "Toroyan").

By appropriate calculations, it was shown that the series corresponding to the dynamics of the appearance of publications in the cases considered have a multifractal nature. At the same time, the corresponding dependences $f(\alpha)$ that correspond to the series under study (Fig. 53) have different curvature parameters. This fact, which is typical for the ratio of topics and subtopics, indicates, on the one hand, that the series that corresponds to the subtopic is less stable than the series corresponding to the entire topic, and on the other hand, that the considered subtopic is not representative for analysis. stream of publications on general topics.

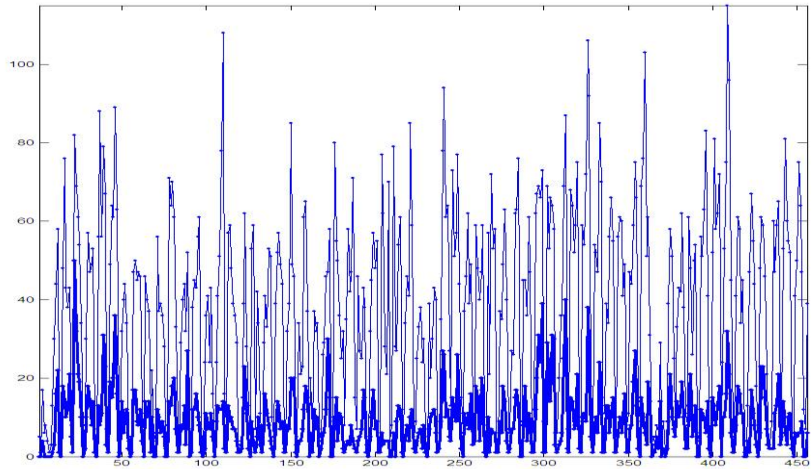


Figure Fig. 52. Diagrams of the intensity of publications on the main (thin connecting line) and refined topics (thick line): abscissa axis - ordinal numbers of days, the y-axis is the number of publications

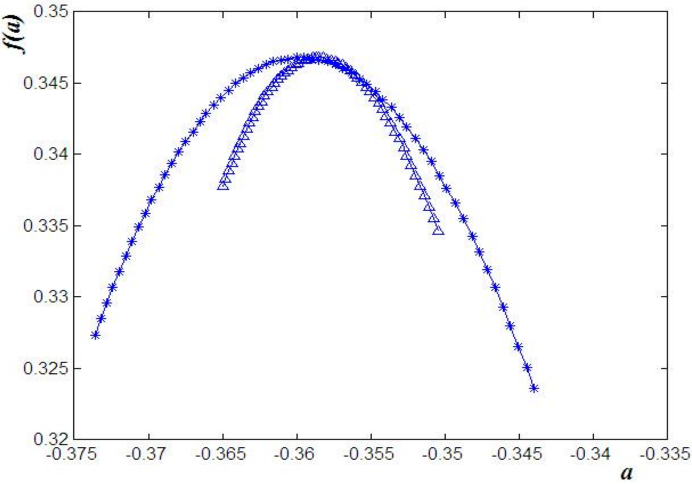


Figure 53. Comparison of multifractal spectra of the studied series by main topic (Δ) and clarified (*) from Singularity Index α

6. VITALITY INFORMATION OPERATIONS

6.1. The concept of "information operations"

information operations has become popular , primarily because information technology plays an ever-increasing role in military operations. At the same time, information operations are defined as "actions aimed at influencing the enemy's information and information systems and protecting one's own information and information systems" [109]. Information operations are seen as combining the basic capabilities of electronic warfare, computer network operations, psychological operations, military operations and security operations in order to influence, destroy, distort information necessary for the enemy to make decisions, as well as protect one's own information.

Information operations cover a whole range of processes carried out in a variety of areas. At the same time, it should be noted that information operations are an essential and traditional component of combat operations. Despite the fact that the formal definition in the documents of the US Department of Defense is focused on the military aspects of information operations, it is quite applicable to almost any area of life.

Below we will consider such information operations that are implemented using information systems. The survivability of these information systems largely determines the survivability of information operations, which are implemented in the form of information impacts on people's consciousness. Information is a reflection of the meaning invested in it, therefore today information has turned from an abstract term into an object, purpose and means of information operations, has become a critical concept in security issues. Former US Secretary of Defense William Cohen March 18 1999 г. stated that "the ability of the army to use information to dominate future battles will give the United States a new key to victory for many years, if not for several generations" [110].

When modeling and conducting information operations, it is necessary to take into account the value of information for decision makers. The value of information includes its timeliness, accuracy and "analyticity". From a practical point of view, the value of information can also be defined as its relevance or applicability, suitability for use. The applicability of information refers to ensuring that decision makers have access to ready-to-use information. The ISO 9241 standard (ISO stands for International Standards Organization) defines applicability in terms of efficiency and satisfaction of the needs of a specified set of users to solve a specified set of tasks in a specific environment. In practice, most of the useful information comes to decision makers from information and analytical systems that provide orientation in the situation and support in making decisions. According to the US War Department's Information Operations Field Manual (FM 100-6), "situational orientation means a combination of a clear understanding of the disposition of friendly and enemy forces with an assessment of the situation and intentions on the part of the command."

Information operations are carried out in a certain social environment, therefore, for their successful implementation, it is necessary to adapt to this environment, to overcome a certain barrier of not very strong attention to information impact. This barrier arises due to the so-called immune system of the environment, which may not miss information impacts if it is powerful enough and/or has already learned to defend itself from such impacts. Preparatory actions for conducting information operations may include the creation of an "immunodeficiency" of the social environment by influencing through the information space, for example, with the help of materials in the media. Very often, information influences use the mechanisms of "viral marketing", for example, in the form of rumors, when sensationally presented disinformation spreads at great speed. It is the immune system that counteracts such information operations. Very often , with the immune system of society , the state is identified, designed to ensure the security of this society , i.e. _ in the presence of a strong state apparatus, the probability of success of antisocial information operations is
venously decreases. The reader knows perfectly well how such informational processes were counteracted in totalitarian states. In a democratic society , of course, totalitarian methods are not applicable. In this case, immunity is achieved through "learning", i.e. a democratic society must go through many informational attacks, influences, influences of stereotypes in order to develop the necessary immunity.

The level of readiness for information operations today is considered a key success factor for any social procedure, campaign.

A special purpose in carrying out information operations is the information and analytical systems of the subject of influence. By influencing such systems, decision makers in the opposing camp can be forced to draw inappropriate conclusions, and the required social process will change trajectory in the direction desired by the influencing party (Figure 54).

In this case, direct information impacts may include the placement in the information space of documents compromising the opposite side, advertising (including hidden) of one's advantages, distorted data about the external environment , distorted information about intentions , etc..

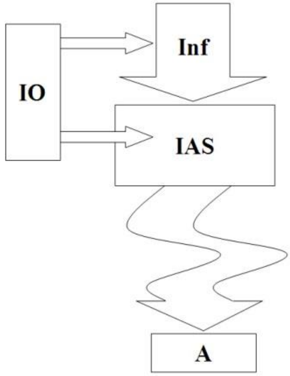


Figure. 54. 54. Impact on the information and analytical system of the enemy:
Inf - information space;
IAS - information and analytical system;
A - system subscriber (decision maker);
IO - information impacts

At the same time, information operations have the following main features:

- information operations is an interdisciplinary set of methods and technologies in such areas as computer science, sociology, psychology, international relations, communications, military science;
- there are still no standards for conducting information operations;
- not only defense departments, but also many government and commercial organizations are interested in the development of information operations technologies;
- the task of forming a scientific approach to information operations is urgent and relevant.

When conducting information operations, it is essential to identify the content (knowledge) invested in information, taking into account a wide variety of aspects - social, political, religious, historical, economic, psychological, mental, cultural, inherent in various strata of society. Therefore, at present it makes sense to consider information operations more broadly, as operations based on knowledge (*English - Know-ledge Operations)* [111].

6.2. Information operations as social procedures

Social procedures and processes tend to be difficult to evaluate and model because their outcomes are psychological and sociological rather than physical. It is this fact that also determines the problematic nature of predicting the results of modeling information operations. In addition, experimentation with information influences in the framework of information operations is more complex and dangerous than in the simulation of physical processes. Actions to be effective in influencing adversary decision-making processes sometimes need to be taken for a long time before they take effect.

One of the main components of information operations is social influence, covering the whole variety of influence processes. Significant changes in people's beliefs or attitudes towards some problem or phenomenon are expected to lead to a change in behavior associated with that problem.

In 1948, Lasswell [112] developed a communication transmission model consisting of five components:

- source - a person who influences or convinces other persons;
- message - with the help of which the source tries to convince the target;
- target - a person whom the source is trying to influence;
- channel — message delivery method;
- impact - the reaction of the target to the message.

Although Lasswell was primarily interested in mass communication, his information transfer model can be applied to interpersonal communication such as the Shannon-Weaver and Osgood-Schramm circular models, which involve feedback loops in the communication process, stating that communication is a circular rather than a linear process [113, 114].

Modeling the objective factors of social influence requires interdisciplinary approaches related to computer science, marketing, political science, and social psychology. The most famous models of the formation of public opinion and social influence are based on the theory of Latane (Latane) dynamic social impact [115, 116], developed by such authors as Nowak, Szamrej, Latane [117], Lewenstein, Nowak and Latane [118], Kacperski and Holyst [119], Sobkowicz [120, 121].

Trying to substantiate the mechanism of the social influence of messages, Latain [115] emphasized the importance of three features of the source-target relationship:

- power - social power, probability or level of influence on individuals;
- immediacy - physical or psychological distance between individuals;
- number of sources - the number of sources tending to the goal.

The current state of information operations modeling is characterized by a number of open problems, the main of which relate to understanding the concepts of information influence and impact.

The universal characteristics of objects are their state and the possibility of influencing other objects. The realization of the possibility of influence requires certain conditions, which are usually called its influence. At the same time, an object that can exercise its will is called a subject, and control is usually called an impact in relation to the object of influence, applied for a specific purpose.

When an individual is the target of influence from one or more sources, dynamic social impact theory states that the level of social influence on an individual can be represented by the following equation, which is the basis of the so-called person -centered model [122]:

$$I_i = -S_i\beta - \sum_{j=1,\, j\neq i}^{N} \frac{S_j O_j O_i}{d_{i,j}^{\alpha}},$$

Where I_i - the magnitude (quantity) of social pressure exerted on the individual i, ($-\infty < I_i < \infty$). O_i represents the individual's opinion (± 1) on the current issue, where +1 and -1 represent support or objection to the issue, respectively. S_i represents the power of the individual i or influence ($S_i > 0$), β is the individual's resistance to change ($\beta > 0$), $d_{i,j}^{\alpha}$ *is* the distance between individuals i And j ($d_{i,j}^{\alpha} \geq 1$), α is the distance reduction indicator ($\alpha \geq 2$), N is the total number of agents (individuals that make up the community). Meaning β, the tendency to hold one's own mind or resist change, determines that individuals within the model may require more or less social pressure to change their minds. Larger levels of value α correspond to the effect of increasing distance between source and target, which affects the amount of social pressure on the target.

On the basis of the terms introduced, the concept of the "information field of an object" is formulated [123] , and its characteristics are described. This makes it possible to define the information impact as an impact on the information field of the object. Exploring the information fields of objects and subjects of social systems, one can determine informational influences and controls. At the same time, information can be considered both as an object and as a means of influence. The use of information as a means of influence requires in the management process to prepare data, produce relevant information, and only then implement the created information in the form of impact (influence).

6.3. Information influence

One of the main methods of conducting information operations is the information influence exerted for the purpose of information management. In this case, information management is understood as a control mechanism when the control action is implicit, indirect informational in nature and the control object is given a certain information picture, under the influence of which it forms its line of behavior. Thus, information management is a method of influence that encourages people to behave in an orderly manner, to perform the required actions.

In accordance with [123, 124], it is expedient to decompose the process of information influence of one object on others into the following stages:

— generation by the source of influence of data, information elements and information sets;

— transmission of information by a source of influence;

— receiving information by the recipient;

— generation of a set of data, information elements and new sets of the object of influence;

- appropriate active actions of the object of influence.

Information impacts on elements of systems can be classified according to such features as sources of occurrence, duration of exposure, nature of occurrence, etc.

To select specific ways to implement information management, it is necessary to specify the tasks solved with the help of information impact, analyze the process of forming information operations and develop criteria for their evaluation. Information management is considered as a process covering the following three interrelated areas:

— management of data exchange between the real world and the virtual world of the subject of influence;

— management of the virtual world of subjects of influence, decision-making mechanisms;

— managing the process of transforming decisions into actions by the subject of influence in the real world.

Information impact can be of two main types:

1. Change in the required direction of the data that the information and analytical system of the object of influence uses when making decisions.

2. Direct influence on the decision-making process of the target, for example, on decision-making procedures or individual decision makers.

The most important for information operations is the environment, the state of the objects of information impact, their mutual influence. In particular, if some electoral field is chosen as the objects of information operations, then it is important to take into account all electoral populations included in this field, which represent supporters (or opponents) of certain political forces. Despite the fact that some models will be considered in the future, in which the homogeneity of the environment is explicitly postulated, in the general case, in relation to information operations, the environment may consist of areas:

- dominant perception;
- hypersensitivity;
- indifference to the relevant informational influences.

6.4. Stages of information operations

Let us dwell separately on the stages of information operations. Obviously, there is no single "standard" plan for conducting both offensive and defensive information operations. One can only consider an approximate sequence of actions obtained by generalizing some already implemented information operations during their implementation.

In practice, an information operation as a process of information influence on mass consciousness is usually implemented as follows: as a result of preliminary intelligence, a plan is developed for the next stage - operational control and appropriate operational intelligence measures are outlined, which are an approximate model of the solution, after which operational control of the enemy is implemented. At the stage of operational intelligence, the level of deviation of the original model from reality is determined, and if it is insignificant, then the original plan is implemented. Otherwise, a new plan of operational command and control of the enemy is being built. The cycle is then repeated until operational intelligence confirms the model. In this case, the final decision is made with a certain operational risk.

Thus, the process of information impact covers such main stages [125] (Fig. 55) as preliminary reconnaissance (PI) (identification of the current situation, enemy state (Op), determination of control tendencies), enemy control (M) (information impact on the enemy in order to convey to him information corresponding to our plan), operational intelligence (OI) (checking the results of reflexive control), operational control (OM) - the actions of our own forces to achieve the required goal.

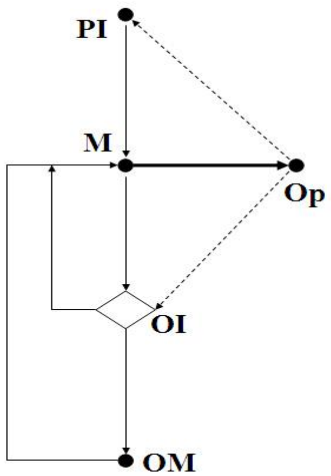


Figure 55. Main stages of information operations

When planning or modeling social processes, in particular information operations, it is always necessary to take into account that the general behavior of social systems cannot be determined using only refined mathematical methods.
sky models. This is the main
at once due to the fact that such processes are largely dependent on socio-psychological factors.

There are two main types of information operations - offensive and defensive. However, in practice, most information operations are mixed. In addition, most information operations procedures are both offensive and defensive. Each of the types of information operations, including the above main stages, implies some features and clarifications.

A feature of offensive information operations (information attacks) is that the objects of influence of such operations are determined and planning is based on sufficiently accurate information about these objects. An information attack most often requires finding or creating an informational occasion (for defensive informational operations, the reason may be the enemy's information attack itself), the promotion of this occasion, i.e. propaganda (as opposed to counter-propaganda measures in defensive information operations), as well as the need to take measures to prevent information counteraction.

Thus, the plan of a typical information operation includes such stages as evaluation, planning, execution, and the final phase, which coincide at the top level for both types of information operations. Let's give a more detailed list of components of information operations.

In offensive information operations, the following main phases can be distinguished:

1. Assessment of the need for an operation:

1) definition of the goal, forecast of achievability, degree of influence;

2) collection of information.

2. Planning.

3. Execution of information impact:

1) finding or creating an information occasion;

2) promotion of an informational occasion (propaganda);

3) operational intelligence;

4) impact assessment;

5) an obstacle to information counteraction;

6) correction of information impact.

4. Final phase:

1) an analysis of efficiency;

2) the use of positive results of information impact;

3) counteraction to negative results.

Typical defensive information information covers the following main stages:

1. Rating:

1) analysis of possible vulnerabilities (goals);

2) collection of information about possible transactions;

3) identification of possible "customers" of information impacts:

– *determination of areas of common interest of the object and potential “customers”;*

- *ranking potential customers according to their interests.*

2. Planning:

1) strategic planning of a defensive operation (explicit or implicit):

- *definition of criteria for information impacts;*

- *modeling of information impacts taking into account: object links; impact dynamics; "special" (critical) points of influence;*

— *forecasting the next steps;*

— *calculation of consequences.*

2) tactical planning of counter-operations.

3. Execution - a reflection of the information impact:

1) identification and “smoothing” of an information occasion;
2) counter-propaganda;
3) operational intelligence;
4) assessment of the information environment;
5) adjustment of information counteraction.
4. Final phase:
1) performance analysis;
2) the use of positive results of information impact;
3) counteraction to negative results.

The operational management of information operations using information and analytical systems can be illustrated using the diagram shown in fig. 56.

In accordance with the above diagram, information from the real world (R) enters the information space, in particular, to the media (I) or directly to experts (E), also through the media. From experts or directly from the information space (for example, using content monitoring tools), information enters the information and analytical system (IAS). The information-analytical system transmits data to decision makers (P) that determine the measures of information impact on the information space and directly on real world objects (people, environment, computer systems, etc.).

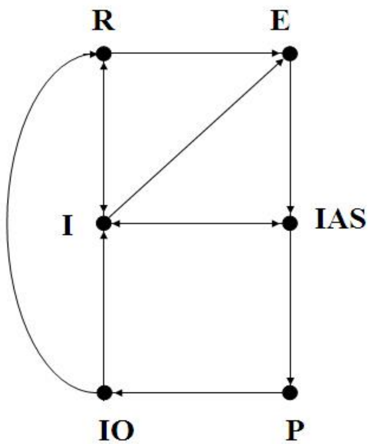


Figure 56. Diagram of operational management using information and analytical systems

6.5. Simulation Features information operations

Modeling can be seen as one of the ways to solve problems that arise in the real world, in particular, when planning and conducting information operations. Most often, simulation is used in cases where experiments with real objects are impossible or too costly. Modeling covers mapping a real problem into an abstract world, learning, analyzing and optimizing the model , and mapping the optimal solution back to the real world.

When modeling, there are two alternative approaches - analytical and simulation modeling. Ideal analytical models allow a rigorous analytical solution, or at least a statement, for example, in the form of systems of differential equations. However, analytical solutions are not always achievable. Therefore, especially in recent times, and especially in solving problems from the field of social dynamics, simulation modeling methods are increasingly being used (*English - Simulation Modeling)*. Simulation is a more powerful and almost indispensable tool for analyzing social procedures. The simulation model can be viewed as a set of rules that determine the future state of the system based on the current one. In this case, the modeling process consists in observing the evolution of the given rules of the system in time, and, accordingly, assessing the adequacy of the model, when possible.

The most promising direction of modeling information operations is the mathematical description of the self-organization of the environment for the perception and dissemination of information, taking into account the current conditions. Self-organizing environments, for which there is no central control mechanism, and development occurs due to many local interactions, are studied by the theory of complex systems. This theory covers such branches of knowledge as nonlinear physics, thermodynamics of nonequilibrium processes, and the theory of dynamical systems. The interactions between individual elements of complex systems determine the occurrence of complex behavior in the absence of centralized control. To study such behavior, the most modern methods are used, covered by the interdisciplinary basis of modern methodology - the concept of complexity. Currently, the theoretical and technological foundations of this concept include the theory of deterministic chaos and complex networks, synergetics, fractal and wave (wavelet) analysis, multi-agent modeling, the theory of self-organized criticality (studying the dynamic development to a critical state, characterized by strong space-time fluctuations, without external control [126]), the theory of percolation (*English - Percolation - flow)*, etc. procedures (information operations , of course , belong to those) involves conducting you - numerical experiments, since most often there are significant limitations that make it difficult to conduct "field" natural experiments.

When modeling information operations, a computational experiment makes it possible to reduce the operations of clarifying constraints, selecting initial data, choosing the rules for the functioning of model components, etc. In this case, it becomes possible to take into account cases that are difficult to implement in practice, using real data only to identify the parameters of the mathematical model. At the same time, mathematical modeling has its limitations; the real world turns out to be difficult to model with a sufficient level of detail and accuracy, i.e. more or less reliable mathematical models are so complex and multi-parametric that they cannot be analyzed and evaluated by exact methods.

It is possible to work out mathematical models when planning information operations only in the process of modeling specific procedures, constantly comparing them with reality.

The expressed purpose of the information operations assessment methodology is to provide a timely and accurate analysis of possible discrepancies between the planned operation and the actual impact. When significant differences are found that affect the probabilities of success of the operation, the analytical system should report this to decision makers in order to correct current plans and decisions. At the same time, when planning information operations, it is impossible to act by trial and error, therefore, it is necessary to develop methods that allow generalizing retrospective data and, on their basis, to check the adequacy of models.

Successful information operations models are based on synergistic approaches. Indeed, society is a complex system, each component of which is characterized by many features, has many degrees of freedom. At the same time, an important property of this system is self-organization, which is the result of the interaction of such components as randomness, repetition, positive and negative feedback [127].

A feature of the mathematical modeling of information operations should be considered the comparative simplicity of interpretation
results. Such concepts as “the size of the electorate”, “political weight”, etc., are perceived on an intuitive level - not even without getting to know the exact (as far as they are possible here) definitions. And this makes it possible to make such an analysis of current situations a subject of wide discussion.

Due to the fact that some solutions are unstable with respect to their parameters, the values of such parameters must be determined with high accuracy. This requires a set of methods based not only on the processing of large volumes of statistical data, but also on versatile sociological research.

At present, the statement of the problem, which consists in using mathematical models to predict possible scenarios of the dynamics of social processes at a qualitative level, looks realistic. In this formulation, dynamics modeling occupies, as it were, an intermediate level between what is stated here and accurate forecasting. Nevertheless, it will be necessary to choose the values of the parameters that would, in some reasonable approximation, correspond to the situation under study, and in most cases the use of relative values turns out to be productive. So, of course, one cannot obtain reliable data on the future development of events, but, most likely, one can form a more or less adequate picture of what can happen and how. And this is not enough.

In order to be successful in this case, individual information impacts must be considered as parts of a single information operation, in the same way that shelling or air attacks can be considered as coordinated parts of a military operation.

6.6. Monitoring and analysis information operations

The modern information space is a unique opportunity to obtain any information on a chosen issue, subject to the availability of appropriate tools, the use of which allows you to analyze the relationship of possible events or events that are already taking place, with the information activity of a certain range of information sources. On the other hand, in a retrospective analysis of any process or phenomenon, certain characteristics of its development are of interest, namely:

- quantitative dynamics inherent in a process or phenomenon, for example, the number of events per unit of time, or the number of messages related to it;
- determination of critical, threshold points that correspond to the quantitative dynamics of the phenomenon;
- determination of manifestations at critical points, for example, identification of the main plots of publications in the media regarding the selected process or phenomenon;
- after identifying the main manifestations of the phenomenon at critical points, these manifestations are ranked, and the dynamics of the development of individual specific manifestations before and after certain critical points is studied;
- statistical, correlation and fractal analysis of the general dynamics and dynamics of individual manifestations is carried out, on the basis of which attempts are made to predict the development of the phenomenon and its individual manifestations.

To study the relationship between real events and publications about them on the Internet, the authors used the InfoStream system, which provides integration and monitoring of network information resources.

The number of web publications per day on any topic, and especially the changes (dynamics) of this value, sometimes allow even small experts in the subject area to draw more or less accurate conclusions.

You can get data on such dynamics, for example, by visiting the sites of news integrators daily (news.yandex.ru, webground.su, uaport.net). Of course, users of professional monitoring systems such as Integrum or InfoStream are in a better position. It is on the basis of the latter system that amazing statistics were obtained on the number of web publications on the subject of influenza epidemics in different periods.

For example, in fig. Figure 57 shows the dynamics of publications in RUNet on the request "bird flu" for the period from mid-2005 to the end of the first half of 2008, obtained using the InfoStream system. Of course, small fluctuations in the number of web publications associated with the weekly cycle can be smoothed out, but still in Fig. 57 three large peak areas are visible, with maxima coming - lasting for December-January for three years. It can be seen how from year to year the topic (even in critical seasons) has lost its relevance. So, there is a periodicity, a decrease in relevance, a bell- shaped form of dynamics in critical seasons.

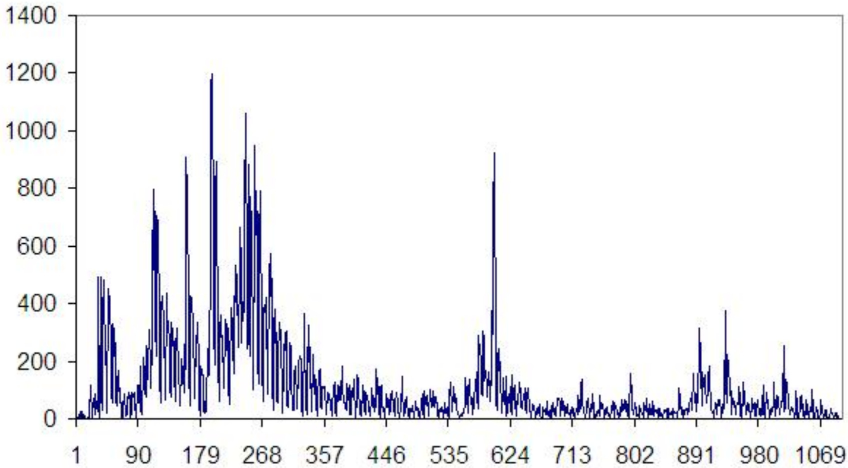


Figure 57. Dynamics of web publications for the query "bird flu"

The dynamics of web publications for the query “swine flu” looks completely different (Fig. 58). The data was obtained from the moment the first messages on this topic appeared in RUNote (second quarter of 2009) until October 2010.

Two sudden peaks can be seen in April 2009 and the end of October 2010. Then the number of web publications decreases sharply - almost in a hyperbolic dependence. The first peak is associated with the first manifestations of the A / H1N1 virus in the world and the release of huge funds around the world (primarily in the United States) to fight it, the other - with the autumn manifestations in all countries of the world, but primarily in Russia and Ukraine. The absolute peak values were more than 3 times higher than the peak values of avian influenza.

Official statistics have already told us that the last epidemic was successfully dealt with, the death rate from A/H1N1 turned out to be lower than from ordinary seasonal influenza. On the face

bursts of web publishing activity that dissipate very quickly. Obviously, reports of swine flu were initially sensational, but then there was no natural information feed. If there had been the Internet in the Middle Ages, then the information about the waves of the plague would probably have had exactly the same character... But the consequences then were different. Indeed, a very convenient platform for supporters of the idea of conspiracy theories and global conspiracies.

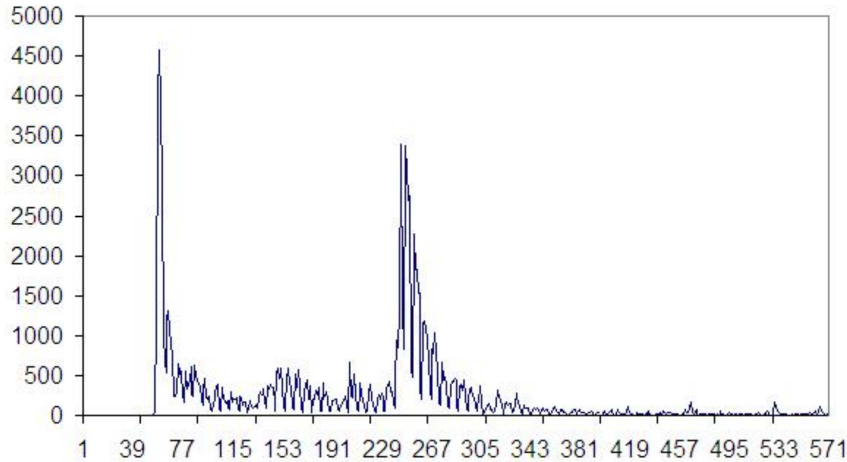
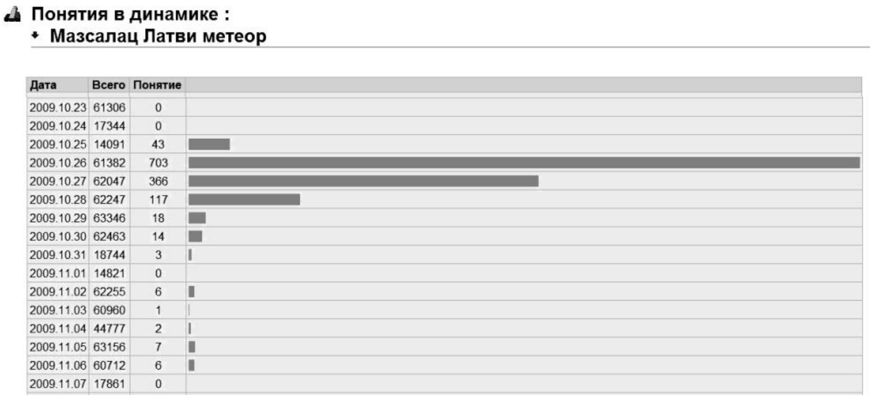


Figure 58. Dynamics of web publications for the query "swine flu"

As we can see, the lack of natural information feed is expressed in a sharp surge in the number of publications, and then also in a rather sharp (hyperbolic or even exponential) decline. Often information operations are accompanied by exactly this behavior of the dynamics of web publications.

Let us give an example of one such information operation carried out at the end of October 2009 by the mobile operator Tele2 from Latvia. It was about a hoax about the fall of a meteorite in the city of Mazsalats. A graph of the dynamics of publications on this topic, obtained using the InfoStream system, is shown in Fig. 59.

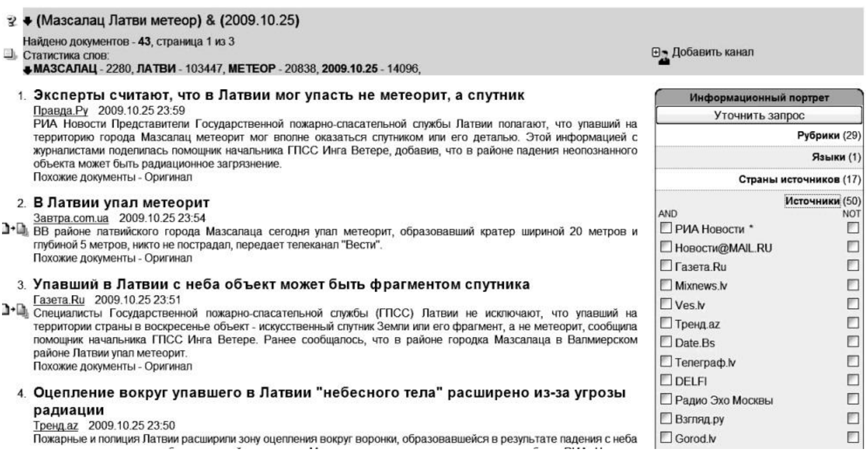


Понятия в динамике :
Мазсалац Латви метеор

Дата	Всего	Понятие
2009.10.23	61306	0
2009.10.24	17344	0
2009.10.25	14091	43
2009.10.26	61382	703
2009.10.27	62047	366
2009.10.28	62247	117
2009.10.29	63346	18
2009.10.30	62463	14
2009.10.31	18744	3
2009.11.01	14821	0
2009.11.02	62255	6
2009.11.03	60960	1
2009.11.04	44777	2
2009.11.05	63156	7
2009.11.06	60712	6
2009.11.07	17861	0

Figure 59. Dynamics of web publications by topic the appearance of a meteorite in Mazsalac

The first publications on October 25 were devoted to the very fact of the appearance of the "meteorite" and attempts to explain its cosmic origin (Fig. 60).



(Мазсалац Латви метеор) & (2009.10.25)
Найдено документов - **43**, страница 1 из 3
Статистика слов:
МАЗСАЛАЦ - 2280, **ЛАТВИ** - 103447, **МЕТЕОР** - 20838, **2009.10.25** - 14096,

Добавить канал

1. **Эксперты считают, что в Латвии мог упасть не метеорит, а спутник**
Правда.Ру 2009.10.25 23:59
РИА Новости Представители Государственной пожарно-спасательной службы Латвии полагают, что упавший на территорию города Мазсалац метеорит мог вполне оказаться спутником или его деталью. Этой информацией с журналистами поделилась помощник начальника ГПСС Инга Ветере, добавив, что в районе падения неопознанного объекта может быть радиационное загрязнение.
Похожие документы - Оригинал

2. **В Латвии упал метеорит**
Завтра.com.ua 2009.10.25 23:54
ВВ районе латвийского города Мазсалаца сегодня упал метеорит, образовавший кратер шириной 20 метров и глубиной 5 метров, никто не пострадал, передает телеканал "Вести".
Похожие документы - Оригинал

3. **Упавший в Латвии с неба объект может быть фрагментом спутника**
Газета.Ru 2009.10.25 23:51
Специалисты Государственной пожарно-спасательной службы (ГПСС) Латвии не исключают, что упавший на территории страны в воскресенье объект - искусственный спутник Земли или его фрагмент, а не метеорит, сообщила помощник начальника ГПСС Инга Ветере. Ранее сообщалось, что в районе городка Мазсалаца в Валмиерском районе Латвии упал метеорит.
Похожие документы - Оригинал

4. **Оцепление вокруг упавшего в Латвии "небесного тела" расширено из-за угрозы радиации**
Тренд.az 2009.10.25 23:50
Пожарные и полиция Латвии расширили зону оцепления вокруг воронки, образовавшейся в результате падения с неба

Информационный портрет
Уточнить запрос
Рубрики (29)
Языки (1)
Страны источников (17)
Источники (50)
AND NOT
РИА Новости *
Новости@MAIL.RU
Газета.Ru
Mixnews.lv
Ves.lv
Тренд.az
Date.Bs
Телеграф.lv
DELFI
Радио Эхо Москвы
Взгляд.ру
Gorod.lv

Figure 60. Publications at the beginning of the studied news story

Apparently, contrary to the expectations of the hoaxers, their exposure received the greatest response in the web space, the information operation went out of control. control of their authors (Fig. 61).

7. **Упавший в Латвии метеорит был размером с кулак, считают ученые**
ФАКТNEWS 2009.10.26 23:10
Метеорит, который в воскресенье вечером упал в Латвии, скорее всего, состоял из железа | delfi.lv Неустановленный "небесный объект", упавший в воскресенье вечером на территории латвийского города Мазсалаца Валмиерского района, по размеру был не больше кулака, сказал в понедельник специалист Центра среды, геологии и метеорологии балтийской республики Улдис Нулле, который ночью побывал на месте ЧП.
Похожие документы - Оригинал

8. **Латышский "метеорит" - это фальшивка. В искусственной воронке взорвали селитру, заподозрили эксперты**
ФАКТNEWS 2009.10.26 23:10
Упавший в латвийском Мазсалаце "небесный объект" может оказаться фальшивкой. Как передает "Интерфакс", к такому выводу пришел геолог Зиемельвидземской биосферической резервации Дайнис Озолс, прибывший утром на место падения | Первый канал Упавший в латвийском Мазсалаце "небесный объект" может оказаться фальшивкой.
Похожие документы - Оригинал

9. **Тайна латвийского метеорита раскрыта - Репортаж**
Петербург-Пятый канал 2009.10.26 23:02
В поле у латвийского городка Мазсалаца сегодня нашли неопознанный объект. Местный фермер раструбил на весь мир: этот огненный шар упал с неба, это метеорит или даже военный спутник.
Похожие документы - Оригинал

10. **Латвийский метеорит оказался пиар-акцией оператора мобильной связи**
Завтра.com.ua 2009.10.26 22:54
Сообщение о падении в Латвии метеорита, ставшее одной из главных новостей местных и мировых СМИ, оказалось шуткой. В инсценировке падения "метеорита" признался оператор мобильной связи Tele2.
Похожие документы - Оригинал

Figure 61. Publications during the "peak" period of the life of the news story

And, finally, the tail of the information plot is completely devoted to the measures of punishment for hoaxers (Fig. 62).

6. **Инсценировка падения метеорита в Латвии не является преступлением**
Политсовет 2009.11.26 16:55
Алена Клименко Сегодня, 26 ноября, помощник начальника Управления полиции Видземского района Латвии Илзе Унгуре заявила, что правоохранители не будут возбуждать уголовный процесс в отношении оператора мобильной связи TELE2, который в октябре инсценировал падение метеорита в районе города Мазсалаца. По словам Унгуре, в действиях компании не обнаружили состава преступления. Вместе с тем, представитель полиции не уточнила, возместил ли оператор все
Похожие документы - Оригинал

7. **Автор шутки с падением "метеорита" в Латвии не будет подвергнут наказанию**
Penki.lt 2009.11.26 16:51
Автор шутки с падением 26 октября "метеорита" в латвийском городе Мазсалаца оператор мобильной связи "Теле2" не будет подвергнут уголовному наказанию, поскольку в его действиях полиция не усматривает состава преступления.
Похожие документы - Оригинал

8. **Заказчика "падения метеорита" в Латвии оштрафуют на 50 долларов**
Газета "Известия" 2009.11.26 16:38
Полиция не нашла в действиях автор шутки с падением 26 октября "метеорита" в латвийском городе Мазсалаца - оператора мобильной связи Tele2 - ничего противозаконного, а потому он не будет подвергнут уголовному наказанию.
Похожие документы - Оригинал

Figure 62. Completion of the life cycle of a news story

The subject of the next studied information flow was determined by a request to the InfoStream system regarding “military operations” in the information space of the country:

(information ~ in oyn & ukrai) | (informats~viyn & ukraï).

Documents relevant to the given request can be presented in two languages (Ukrainian and Russian), contain phrases such as "Information War", or "Information War", and also contain the name of our country. The above queries correspond to the concept of "information wars", which is most often used in the web environment as a functional synonym for "information operations". The retrospective period of the study represented the entire year 2008. As a result of the search for the broadest query, 6196 documents were found. Based on the processing of these data, a complete picture of the experimental data was obtained - time series for a given period. On fig. 63 shows a graph of the number of publications on demand by day in 2008.

The presented schedule takes into account weekly fluctuations (on weekends, for example, significantly fewer documents are published on the Internet than on weekdays).

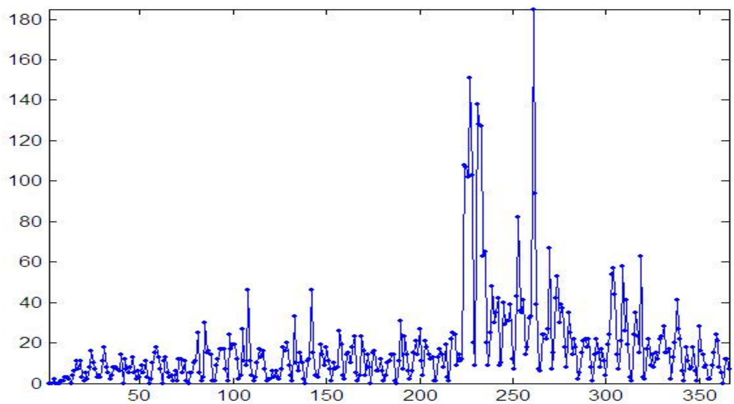


Figure 63. Dynamics of the number of publications on demand by day in 2008 (together 6196 publications)

For a more visual display of trends, such graphs are smoothed using the “moving average” method with an observation window of 7 days. On fig. 64 shows a smoothed graph corresponding to the above dynamics. In particular, it can be seen that around the 220th day of the year, the total number of messages on the subject of information wars increased dramatically.

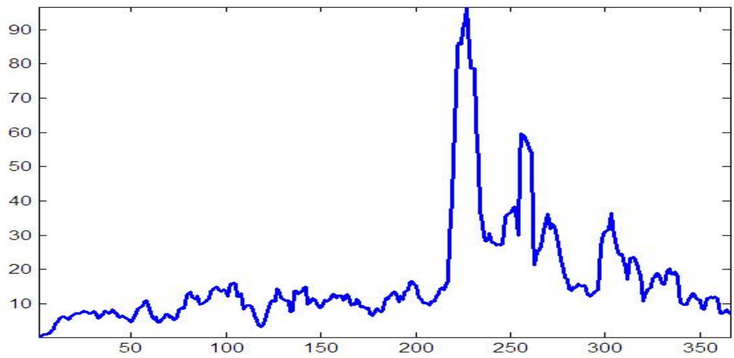
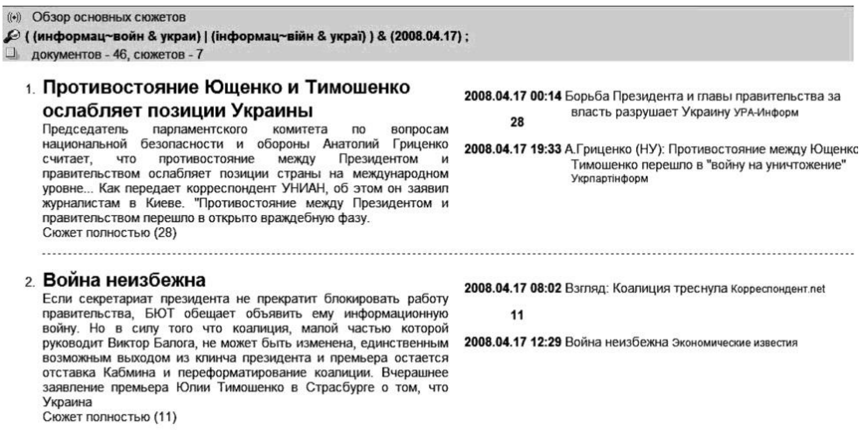


Figure 64. Smoothed graph of the number of publications by day in 2008

On fig. Figures 65-67 show the content of messages summarized by the InfoStream system, which refer to some peak values of the intensity of publications during 2008.

The task of studying the statistical properties of network documentary arrays [65] is multifaceted and allows the active use of modern methods, including fractal analysis methods [128–130], which allow a deeper understanding of the specifics of the subject area.


Обзор основных сюжетов
((информац~войн & украи) | (інформац~війн & украї)) & (2008.04.17) ;
документов - 46, сюжетов - 7

1. **Противостояние Ющенко и Тимошенко ослабляет позиции Украины**
Председатель парламентского комитета по вопросам национальной безопасности и обороны Анатолий Гриценко считает, что противостояние между Президентом и правительством ослабляет позиции страны на международном уровне... Как передает корреспондент УНИАН, об этом он заявил журналистам в Киеве. "Противостояние между Президентом и правительством перешло в открыто враждебную фазу.
Сюжет полностью (28)

2008.04.17 00:14 Борьба Президента и главы правительства за власть разрушает Украину УРА-Информ
28
2008.04.17 19:33 А.Гриценко (НУ): Противостояние между Ющенкс Тимошенко перешло в "войну на уничтожение" Укрпартінформ

2. **Война неизбежна**
Если секретариат президента не прекратит блокировать работу правительства, БЮТ обещает объявить ему информационную войну. Но в силу того что коалиция, малой частью которой руководит Виктор Балога, не может быть изменена, единственным возможным выходом из клинча президента и премьера остается отставка Кабмина и переформатирование коалиции. Вчерашнее заявление премьера Юлии Тимошенко в Страсбурге о том, что Украина
Сюжет полностью (11)

2008.04.17 08:02 Взгляд: Коалиция треснула Корреспондент.net
11
2008.04.17 12:29 Война неизбежна Экономические известия

Figure 65. A fragment of the main stories for April 17, 2008
(dominant theme – performance Prime Minister of Ukraine in Strasbourg)

Обзор основных сюжетов
((информац~войн & украи) | (інформац~війн & украї)) & (2008.08.14) ;
документов - 151, сюжетов - 29

1. **Балтийские пособники кавказской трагедии**
"Южная Осетия - это очень небольшой регион, и я считаю, что ее воссоединение с Грузией - вопрос нескольких месяцев". Это заявление Михаила Саакашвили примечательно тем, что оно сделано не в начале августа 2008-го, а четыре года назад, в интервью эстонской газете "Постимеэс" накануне визита президента Грузии в Латвию, Литву и Эстонию.
Сюжет полностью (42)

2008.08.14 00:00 Волки и овцы Российская газета
42
2008.08.14 21:01 Кавказский Саддам "2000" Еженедельник

---

2. **"Нас никто не остановит при возвращении на базу в Севастополь"**
"Новая газета" (Россия) Конфликт вокруг Черноморского флота. Наш собственный корреспондент - из Крыма Во вторник 12 августа президент Украины Виктор Ющенко вновь стоял на Майдане - на площади Свободы в центре Тбилиси, куда он прибыл с коллегами из Польши, Эстонии, Литвы и Латвии.
Сюжет полностью (36)

2008.08.14 00:19 Исследование: Украина проиграла информационную войну России в освещении конфликта в Грузии "Завтра"
36
2008.08.14 21:22 Новые информационные войны Газета "День"

---

3. **Российские хакеры-"миротворцы" напали на украинский портал**
Портал delo.ua подвергся хакерской атаке из-за сегодняшних публикаций в газете "Дело". Об этом говорится в пресс-релизе киевского издательства "Экономика", поступившем в "Обком": "Сегодня в 11 часов утра по киевскому времени украинский портал delo.ua подвергся DDoS атаке с десятка тысяч компьютеров, расположенных по всему миру.
Сюжет полностью (9)

2008.08.14 16:09 Сайт delo.ua атакували через матеріали про російсько-грузинську інформаційну війну Телекри
9
2008.08.14 21:34 Русские хакеры атаковали украинский сайт Обозреватель

Figure 66. A fragment of the main stories for August 14, 2008 (the dominant topic is the Russian-Georgian military conflict)

Обзор основных сюжетов
((информац~войн & украи) | (інформац~війн & украї)) & (2008.11.14) ;
документов - 63, сюжетов - 12

1. **Украина могла бы эффективнее защищаться от России в информационной войне**
Украинское министерство иностранных дел возмущено действиями посла России в Украине Виктора Черномырдина. Он выступил "соорганизатором откровенно провокационной антиукраинской акции", во время которой планировался показ российского фильма "Искусство предательства". В фильме, который так хотело показать украинцам российское посольство, речь идет о российско-грузинском конфликте и якобы участии в нем украинцев.
Сюжет полностью (34)

2008.11.14 01:09 Посольство России требует от Киева объяснени за "грубость" Росбалт Украина
34
2008.11.14 21:41 Россия ведет информационную войну относител Украины Vlasti.net

---

2. **Росія проводить інформаційну війну щодо України - експерт**
Українське міністерство закордонних справ обурене діями посла Росії в Україні Віктора Черномирдіна. Він виступив "співорганізатором відверто провокаційної антиукраїнської акції", під час якої планувався показ російського фільму "Мистецтво зради". У фільмі, який так хотіло показати українцям російське посольство, йдеться про російсько-грузинський конфлікт та нібито участь у ньому українців.
Сюжет полностью (11)

2008.11.14 10:03 Росія веде проти України інформаційну війну Газе по-українськи
11
2008.11.14 19:02 Росія веде проти України інформаційну війну Рух-I

Figure 67. A fragment of the main stories for November 14, 2008 (the dominant theme is a Russian film "The Art of Betrayal"

On fig. 68 presents ratios R/S for a number of publications by day in 2008, corresponding to the above query. The slope of this straight line corresponds to the Hurst exponent.

As can be seen, the Hurst value for the studied information flows corresponds to ~0.81, which confirms the assumption of self-similarity and iteration of processes in the information space. This means that resonant publications are repeatedly duplicated, retold, and discussed. This also means that the general information tension remains at a high level, as soon as the “trail” of one story on the topic of information operations disappears, it is replaced by a new story, most often, as trends show, more intense.

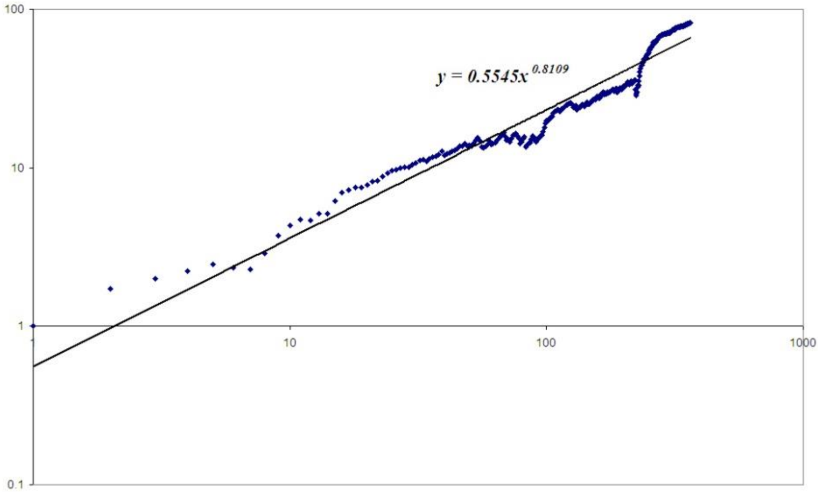
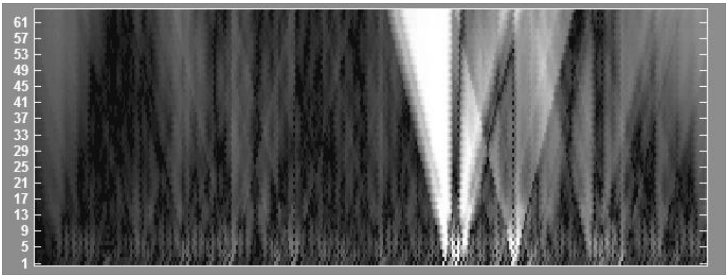


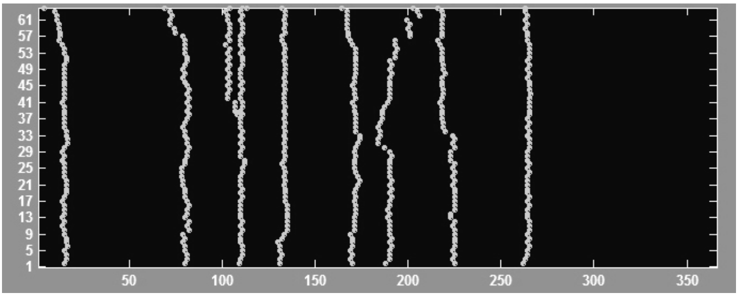
Figure 68. Normalized range indicator on a logarithmic scale for the entire observation period

The most common tools for mathematical modeling and evaluation of observation series also include wavelet analysis [100, 101]. Each of the main factors of the dynamics of the initial process has its own characteristic reflection on the scalogram, while all analytical information is presented in a visual and easy-to-study form. On fig. The scalogram shown in Figure 69 is the result of a continuous wavelet analysis (Haar wavelet) of a time series that corresponds to the process under study.

On the skeleton for most frequencies, not only the 220th day, but also implicit extremes (105th, 130th, 200th days, etc.) are noted. which influence public thought and, ultimately, the information security of both business and the state.



A



b

Figure 69. The result of wavelet analysis (continuous wavelet transform): a - wavelet scalegram; b - lines of local maxima (skeleton)

A common network information attack in the web environment today is carried out as follows: as a rule, a website is created and operates for some time (let's call it the "original source"), while it publishes quite correct information. At hour X, a document appears on his page, usually compromising information on the object of attack, reliable or falsified. Then there is the so-called “laundering of information”. The document is reprinted by Internet publications of two types - those interested in the attack and those who simply do not have enough information to fill their information field. In the case of claims, all reprint publications refer to the "original source" and, as a last resort, at the request / demand of the object of attack, remove information from their websites. The primary source, if necessary, also removes information or is completely eliminated (after which it turns out that it is registered on the Internet to a non-existent person). At the same time, the information has already spread, the task of the original source has been completed, the attack has started.

Everyone is aware of the information campaign directed against Prominvestbank, which began at the end of September 2008 г. With the help of the content monitoring system InfoStream (http://infostream.ua), which scans all the main information websites of Ukraine in real time, the dynamics of publications on the websites of messages that mentioned "Prominvestbank" for three months was determined - September, October and November (Fig. 70). This dynamics testifies to a small number of publications in the first half of September, but then a number of publications began to compromise the chairman of the board V. Matvienko, which caused a relatively small response.

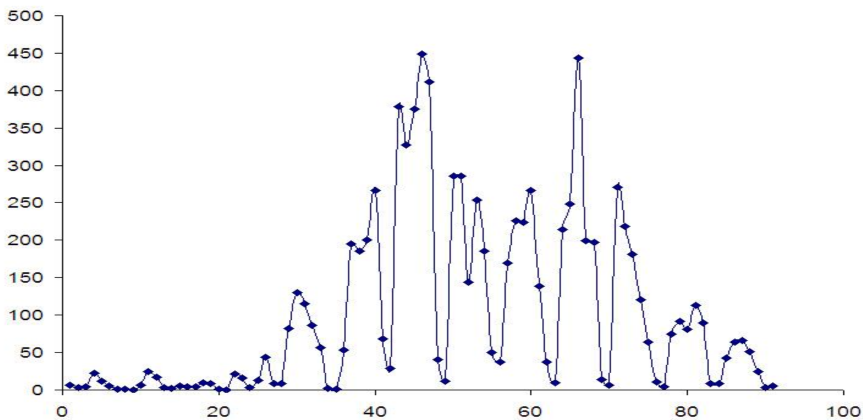


Figure 70. Dynamics of publications on the topic "Prominvestbank" for three months 2008 г.

As it turned out later, these publications were only "artificial preparation". On September 26, the first reports of a possible bank failure appeared (Figure 71), the number of which was quite consistent with an avalanche process, limited only by the number of websites capable of publishing such information. However, this process reached a stable-average level by December 2008 г.

Kramatorsk.info **2008.09.26** 19:35
http://www.kramatorsk.info/?view&62181

В Донбассе вошла в активную фазу атака на Проминвестбанк. ПИБ заявляет, что это атака из-за рубежа

Сегодня в Донецкой области население организованно вышло к проходным **Проминвестбанка.**

Вести о том, что народные массы Донбасса штурмуют отделения ПИБа в Донецке, Авдеевке, Волновахе и пр. населенных пунктах Донецкой области, приходят в "Обком" с середины дня.

Никто из опрошенных нами экспертов не может пока сказать что-либо конкретное по данному поводу, кроме банальных констатаций: ПИБ - серьезный банк, он кредитует промышленный сектор Украины, Донецкое облотделение ПИБа - одно из крупнейших, борьба за него началась еще в середине 90-х годов... Ну а баннеры на киевских дорогах против нынешнего (неизменного) руководства ПИБа во главе с г-ном Матвиенко видели многие автомобилисты и пассажиры столичного транспорта.

"Обком" пока не готов сказать что-то определенное по поводу паники, которая охватила сегодня трудовой Донбасс - хотя сведения для определенных умозаключений, в принципе, имеются. Вместо этого мы предлагаем внимаю вкладчиков сообщение, поступившее от пресс-службы ПИБа:

"**Проминвестбанк** заявляет о стабильной работе, несмотря на дезинформацию в ряде СМИ о якобы приближающемся банкротстве банка.

Проминвестбанк, по оценкам зарубежных экспертов, стабильный банк и занимает в Украине второе место по надежности.

Массовая газетная атака на **Проминвестбанк** организована рейдерскими (бандитскими) группировками зарубежных агентов с участием высокопоставленных чиновников крупных государственных структур, которые по Конституции должны защищать отечественные предприятия и банки. Ложь, шантаж, направленные против банка, преследуют цель вынудить его к продаже иностранцам за комиссионное вознаграждение... Заявляем: банк не продается... **Проминвестбанк** останется украинским!", - говорится в сообщении.

Служба информационной поддержки **Проминвестбанка** также сообщает, что, несмотря на беспокойство вкладчиков, вызванное негативными публикациями о банке, все обязательства перед клиентами и вкладчиками выполняются, а структурные подразделения банка работают в нормальном режиме.

"Обком"

Figure 71. One of the first alarm messages

It cannot be argued that only an information attack via the Internet led the bank to a sad state, but it was the first alarming messages that undermined the confidence of many depositors and forced them to massively withdraw their savings from the bank. On September 30, it was reported that in order to save the bank, the NBU decided to allocate UAH 5 billion in refinancing to Prominvestbank, and on December 5, it was reported that Prominvestbank had a new owner (Fig. 72). After that, the volume of publications about Prominvestbank significantly decreased, which testifies not so much to its recovery, but to the systemic crisis of the banking system of Ukraine, which “dropped” many other credit and banking institutions.

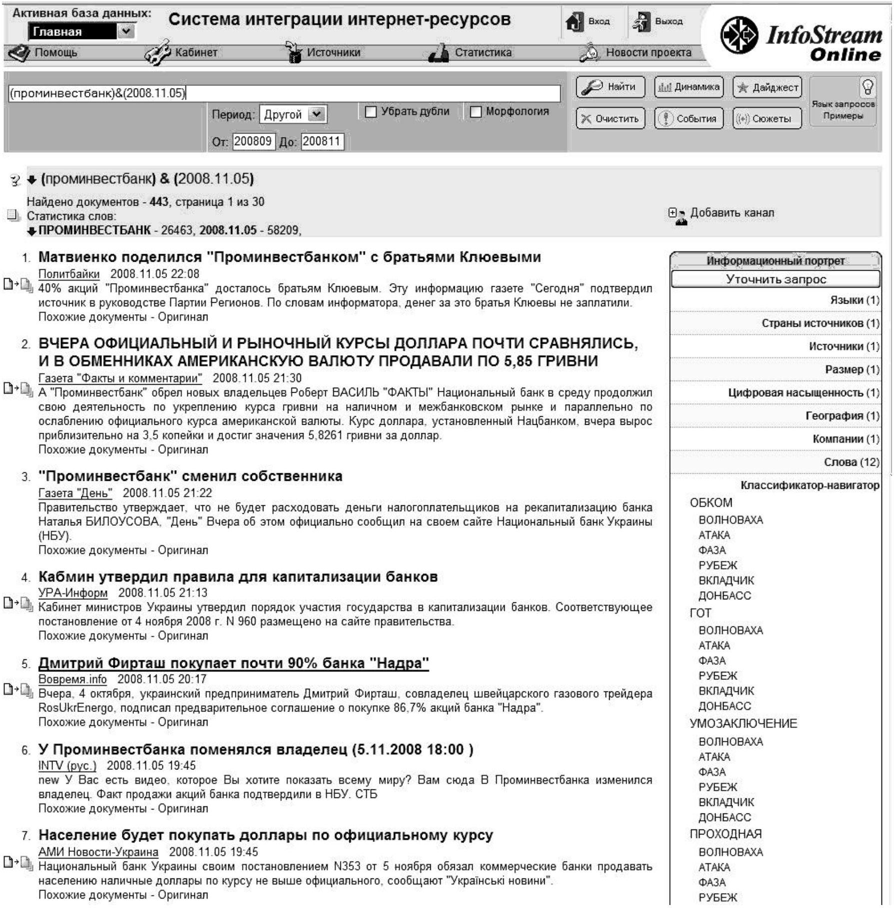


Figure 72. Messages that completed extreme dynamics intensity of publications on the topic "Prominvestbank"

Literally a week after the events described above, another public landmark information attack took place in Ukraine, this time on the insurance market. It was a real information operation against NJSIC "Oranta". In this case, the primary source of compromising material was not a website, but an informational message sent by e-mail to thousands of Internet users. As a result of the use of special techniques, it diverged from the designation of the address of the press service of the object of attack. So, on December 10, 2008, around 11:30 am, an informational message was sent in the form of spam , stating that the Oranta insurance company was declaring bankruptcy. According to preliminary data, the information was scattered to 1000 addresses, naturally, the data got to competitors and the media. The message said that from December 31, 2008, the company ceases to fulfill its obligations to customers.

In connection with the incident, NJSIC "Oranta" appealed to law enforcement agencies with a request to investigate this incident and punish those responsible. What happened with Oranta was very reminiscent of the situation with Prominvestbank, numerous experts agreed with this. After all, both the banking business and the insurance business are based on the trust of customers, which is most easily undermined by information attacks. According to Oleg Spilka, Chairman of the Supervisory Board of NJSIC "Oranta", "This event was deliberately prepared in order to discredit the insurance company and undermine its reputation." Without going into details of the possible targets of the attack (change of owners, struggle for a blocking stake, destruction of the company, etc.), with the help of a retrospective analysis, we will follow the dynamics of publications on the Internet that mentioned NJSIC "Oranta".

On fig. 73 shows the daily dynamics of the number of relevant publications. This chart, among other things, clearly shows the decline in the intensity of publications on this topic in early December 2008 г, which can be perceived as some "calm before the storm".

For the analysis of time series in the framework of the study of the authors, ΔL the - method was used. On fig. 74 shows the scalogram of the dynamics of the process under consideration using the (ΔL-method) for the second half of 2008. Despite separate peaks on the 16th and 55th days of the quarter, the extremum that falls exactly on December 10-12 is of the greatest interest.

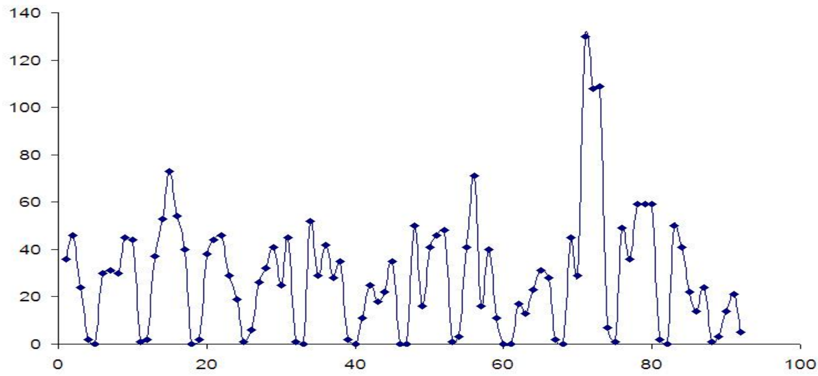


Figure 73. Intensity of publications on the Internet on the topic "Oranta"

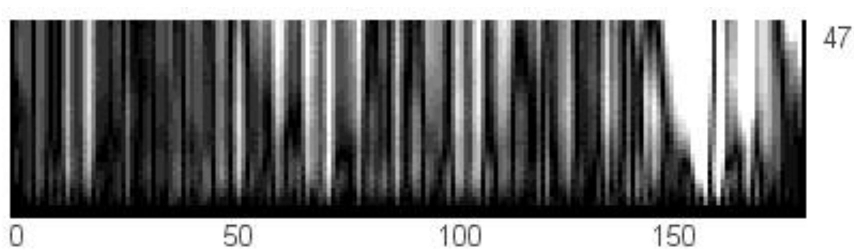


Figure 74. ΔL -diagram of a number of publications on the topic "Oranta"

More detailed statistics of publications on the topic "Oranta" for December 2008 was obtained through the user interface of the InfoStream content monitoring system (Fig. 75).

Let's follow the progress of the information operation, considering the messages published at different time intervals.

On fig. 76 shows a list of publications on the topic "Oranta" during the first hours of the attack. According to Oleg Spilka, within two hours from the beginning of the attack, all the mail servers of NJSIC "Oranta" were disabled, so the denial on the network was delayed.

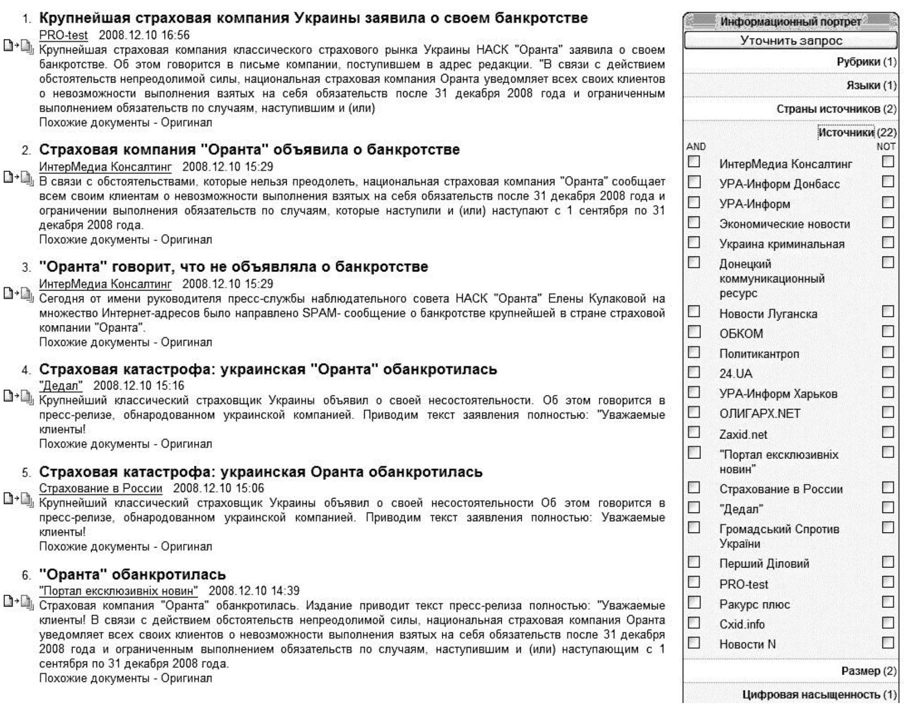
2008.09.14	12567	2
2008.09.15	55451	5
2008.09.16	57084	4
2008.09.17	58559	4
2008.09.18	57273	10
2008.09.19	55868	8
2008.09.20	15395	1
2008.09.21	12126	0
2008.09.22	55948	21
2008.09.23	57232	16
2008.09.24	58203	3
2008.09.25	57520	13
2008.09.26	54628	44
2008.09.27	16331	9
2008.09.28	12572	8
2008.09.29	55595	82
2008.09.30	57436	130
2008.10.01	57117	115
2008.10.02	58419	86
2008.10.03	56392	56
2008.10.04	14930	2
2008.10.05	11291	1
2008.10.06	56672	53
2008.10.07	59109	195
2008.10.08	58352	186
2008.10.09	59923	200
2008.10.10	57726	267
2008.10.11	15692	68
2008.10.12	13165	29
2008.10.13	57178	378
2008.10.14	58234	327
2008.10.15	59271	375
2008.10.16	58517	449
2008.10.17	56452	412
2008.10.18	15300	40
2008.10.19	11404	12
2008.10.20	56814	286
2008.10.21	58055	286
2008.10.22	58192	144
2008.10.23	57697	254

Figure 75. Detail diagram intensity of publications on the topic "Oranta"

At 12:31 pm, a strange “updated” message appears on the Economic News website with a paradoxical last sentence (Fig. 77).

Further, the management of NJSIC "Oranta" published the first denials on the Internet, slowly blaming competitors for what had happened , and then nevertheless recognizing the attack as targeted and beneficial to third parties.

On fig. 78 shows a list of publications devoted to the refutation of the bankruptcy report for the next day (December 11), as well as the most active sources that published these reports. Of undoubted interest of analysts is the comparison of the sources shown in fig. 76 and 78.



1. **Крупнейшая страховая компания Украины заявила о своем банкротстве**
PRO-test 2008.12.10 16:56
Крупнейшая страховая компания классического страхового рынка Украины НАСК "Оранта" заявила о своем банкротстве. Об этом говорится в письме компании, поступившем в адрес редакции. "В связи с действием обстоятельств непреодолимой силы, национальная страховая компания Оранта уведомляет всех своих клиентов о невозможности выполнения взятых на себя обязательств после 31 декабря 2008 года и ограниченным выполнением обязательств по случаям, наступившим и (или)
Похожие документы - Оригинал

2. **Страховая компания "Оранта" объявила о банкротстве**
ИнтерМедиа Консалтинг 2008.12.10 15:29
В связи с обстоятельствами, которые нельзя преодолеть, национальная страховая компания "Оранта" сообщает всем своим клиентам о невозможности выполнения взятых на себя обязательств после 31 декабря 2008 года и ограничении выполнения обязательств по случаям, которые наступили и (или) наступают с 1 сентября по 31 декабря 2008 года.
Похожие документы - Оригинал

3. **"Оранта" говорит, что не объявляла о банкротстве**
ИнтерМедиа Консалтинг 2008.12.10 15:29
Сегодня от имени руководителя пресс-службы наблюдательного совета НАСК "Оранта" Елены Кулаковой на множество Интернет-адресов было направлено SPAM- сообщение о банкротстве крупнейшей в стране страховой компании "Оранта".
Похожие документы - Оригинал

4. **Страховая катастрофа: украинская "Оранта" обанкротилась**
"Дедал" 2008.12.10 15:16
Крупнейший классический страховщик Украины объявил о своей несостоятельности. Об этом говорится в пресс-релизе, обнародованном украинской компанией. Приводим текст заявления полностью: "Уважаемые клиенты!
Похожие документы - Оригинал

5. **Страховая катастрофа: украинская Оранта обанкротилась**
Страхование в России 2008.12.10 15:06
Крупнейший классический страховщик Украины объявил о своей несостоятельности Об этом говорится в пресс-релизе, обнародованном украинской компанией. Приводим текст заявления полностью: Уважаемые клиенты!
Похожие документы - Оригинал

6. **"Оранта" обанкротилась**
"Портал ексклюзивних новин" 2008.12.10 14:39
Страховая компания "Оранта" обанкротилась. Издание приводит текст пресс-релиза полностью: "Уважаемые клиенты! В связи с действием обстоятельств непреодолимой силы, национальная страховая компания Оранта уведомляет всех своих клиентов о невозможности выполнения взятых на себя обязательств после 31 декабря 2008 года и ограниченным выполнением обязательств по случаям, наступившим и (или) наступающим с 1 сентября по 31 декабря 2008 года.
Похожие документы - Оригинал

Информационный портрет
Уточнить запрос
Рубрики (1)
Языки (1)
Страны источников (2)
Источники (22)
AND NOT
ИнтерМедиа Консалтинг
УРА-Информ Донбасс
УРА-Информ
Экономические новости
Украина криминальная
Донецкий коммуникационный ресурс
Новости Луганска
ОБКОМ
Политикантроп
24.UA
УРА-Информ Харьков
ОЛИГАРХ.NET
Zaxid.net
"Портал ексклюзивних новин"
Страхование в России
"Дедал"
Громадський Спротив України
Перший Діловий
PRO-test
Ракурс плюс
Cxid.info
Новости N
Размер (2)
Цифровая насыщенность (1)

Figure 76. The first hours of the attack. The most "operational" sources

The further decline in publications on the topic of NJSIC "Oranta" and its return to the normal "average" level indicates that the company, with its careful and precise actions, was able to successfully resist the information operation.

Экономические новости **2008.12.10** 12:31
http://economic-ua.com/articles/46840

Страховая компания "Оранта" стала банкротом (обновлено)
В Интернете появились сообщения о том, что страховая компания "**Оранта**" стала **банкротом.**

"Уважаемые клиенты!
В связи с действием обстоятельств непреодолимой силы, национальная страховая компания **Оранта** уведомляет всех своих клиентов о невозможности выполнения взятых на себя обязательств после 31 декабря 2008 года и ограниченным выполнением обязательств по случаям, наступившим и (или) наступающим с 1 сентября по 31 декабря 2008 года.
В связи с начатой процедурой **банкротства** действие всех страховых полисов ограничивается сроком до 31 декабря 2008 года, вне зависимости от даты, указанной в договоре.
Страховые возмещения по случаям, наступившим с 1 сентября по 1 декабря 2008 года, будут выплачены в период от одного до трех лет, от даты судебного решения о **банкротстве.** С 1 января 2008 года ответственность по полисам НАСК "**Оранта**" будет переложена на ряд партнерских страховых компаний. Список партнеров будет опубликован на нашем сайте.
Клиентам, у которых срок действия договоров заканчивается позже 31-го декабря, необходимо прибыть в ближайшее отделение компании и перезаключить договор страхования с нашими партнерами. До 31-го декабря на перезаключение договоров по абсолютно всем видам страхования нашими партнерами предоставляются скидки". На официальном сайте компании, данная информация не подтвердилась.

Как сообщили "ЭН" в самом НАСК "**Оранта**", это не правдивая информация.

Figure 77. Refutation?

1. **"Оранта" - не банкрот**
   Бердянские ведомости 2008.12.11 19:49
   Смирнова Екатерина В среду, 10 декабря, в интернет-изданиях появилась информация о том, что крупнейшая страховая компания классического страхового рынка Украины НАСК "Оранта" заявила о своем банкротстве и уведомила своих клиентов о невозможности выполнения взятых на себя обязательств после 31 декабря 2008 года.
   Похожие документы - Оригинал
2. **Официальный пресс-релиз НАСК Оранта по поводу мнимого банкротства**
   News.kompass.ua 2008.12.11 17:35
   10 декабря, в 11.30 от имени руководителя пресс-службы наблюдательного совета НАСК Оранта Алены Кулаковой на множество Интернет -адресов было направлено СПАМ-сообщение о банкротстве крупнейшей в стране страховой компании Оранта.
   Похожие документы - Оригинал
3. **Страховая компания "Оранта" подверглась массированной информационной атаке**
   ПростоБанк.ua 2008.12.11 17:07
   Многие интернет-пользователи получили по электронной почте сообщение, в котором якобы от имени руководителя пресс-службы "Оранты" говорилось о том, что страховщик начал процедуру банкротства и с 31 декабря 2008 года прекращает выполнять перед своими клиентами все взятые на себя обязательства.
   Похожие документы - Оригинал
4. **"Оранта" исключает причастность конкурентов к информации о якобы банкротстве компании**
   УНІАН 2008.12.11 16:46
   Руководство национальной акционерной страховой компании (НАСК) "Оранта" исключает причастность конкурентов к распространению через спам-рассылку информации о якобы банкротстве компании.
   Похожие документы - Оригинал
5. **"Оранта" намерена привлечь к ответственности распространителей лже-информации**
   УНІАН 2008.12.11 16:46
   Руководство национальной акционерной страховой компании (НАСК) "Оранта" обратилось в Генеральную прокуратуру, Службу безопасности Украины, а также к премьер-министру Украины с просьбой о расследовании инцидента и привлечении к ответственности распространителей информации о якобы банкротстве компании.
   Похожие документы - Оригинал
6. **"Оранта" просит Генпрокуратуру найти автора спама о банкротстве**
   "Мир домов" 2008.12.11 15:57
   Утром 10 декабря от имени руководителя пресс-службы СК "Оранта" было Елены Кулаковой было отправлено на множество интернет-адресов спам-сообщение о банкротстве страховой компании.
   Похожие документы - Оригинал

Информационный портрет
Уточнить запрос
Рубрики (6)
Языки (2)
Страны источников (3)
Источники (19)
AND NOT
МЕТА - Украина. Пресс-релизы
СТРАХНАДЗОР
PR - это жизнь
TRISTAR.com.ua
Економічна правда
UkrBiz.net
УНІАН
Одеські вісті
forINSURER.com
"Страхование сегодня"
B2Blogger.com
"Коммерсантъ-Украина"
Prboom
"Вікна" Калущини
Fin.org.ua
Міністерство економіки України
НАСК "Оранта"
Страхование Украины
Украина деловая
Размер (2)
Цифровая насыщенность (3)
Тональность (2)
География (25)

Figure 78. Rebuttal messages

The considered practical examples made it possible to develop some general methodology for conducting a defensive information operation using a web resource content monitoring system. Suppose the object of an aggressive information operation is the company "ABV". The following 12 countermeasures are suggested:

1. Collection of information with publications in "foreign" (not related to "ABV", unaffiliated) media about the company.

2. Construction of a graph - the dynamics of the appearance of messages about the company "ABV" in the online media.

3. Analysis of dynamics with a retrospective of 6–12 months using the methods of time series analysis. After that, the content of publications is analyzed at threshold points, the moments, duration, frequency of impact are determined, and the binding of the moments of impact to other events from the area of interest of the object.

4. Determining the sources that publish the largest number of negative (publications with a negative tone) about the company "ABV".

5. Determination of the "primary sources" of publications in the media - those sources that were the first to publish negative information.

6. Identification of probable "customers" - owners or persons influencing the publishing policy of individual media.

7. Determining the areas of common interest of the ABV company and potential "customers" (by identifying common information characteristics - intersections of the "information portraits" of the InfoStream system built for the facility and the "customer"), ranking potential "customers" according to their interests.

8. Determining the criteria for information impacts based on the most rated interests.

9. Modeling information impacts, for which the “customer” connections are found - the persons and organizations most associated with it, the dynamics of the impact from the customer is analyzed and a forecast of this dynamics is built, the content of publications is analyzed at the threshold points of the dynamics curve - critical impact points are determined.

10. Further impact steps are predicted by analyzing similar publishing trends for other companies in the InfoStream retrospective database.

11. Taking into account the realities and publications from the retrospective database, the likely consequences are assessed.

12. Informational (and not only) opposition is being organized. Examples of publications in the context of resistance are in the retrospective database.

It is content monitoring systems that are best suited for the operational analysis of the information situation for three reasons. Firstly, they provide efficiency that search engines cannot provide (the time for indexing network content even by the best of them ranges from several days to several weeks); secondly, specialized content monitoring systems provide completeness both in terms of sources and presentation of source materials, while conventional news aggregators do not always provide the necessary completeness; and thirdly, content monitoring systems contain the necessary analytical tools that can provide the user with information about the intensity of publications on a given topic in the required period of time.

In terms of prevention of information operations, one should carefully monitor the dynamics of publications about the target company, if possible, taking into account the tone of these publications, use available analytical tools, for example, wavelet analysis. At the same time, one should be guided by possible models of information attacks, for example, if this model covers the phases: “background publications” - “calm” - “artillery preparation” - “calm” - “attack” (Fig. 79), then already for the first three components it is possible to predict future events with a high probability.

The above plan is obviously the ideal one, focused solely on web resource content monitoring data.

At the same time, the pattern found by the authors found its continuation and mathematical justification [131]. The above phases of information attacks, displayed as local and global maxima on the time series, made it possible to construct a template to which the Kunchenko polynomials were applied [132].

In the above model, the extrema correspond to the phases of "attempt" and "attack". In this case, the second maximum will be greater than the first. The difference between these maximums can be very significant, the “attack” phase will correspond to the interval with the highest density of messages published throughout the day.

Figure 79. Typical behavior of series of intensity of thematic publications

The selected patterns of behavior of the intensity series of thematic publications can be considered as patterns (samples) of functional dependence. These templates can be taken as a single basic element of some linear space, i.e. as a generating element e for modeling using Kunchenko polynomials.

Whereas a linear combination of linearly independent transformations $f_1(e), f_2(e),..., f_n(e)$ of the corresponding generating element can be constructed by an P_n approximation polynomial n of -th order to a part of the output signal $f_s(e)$:

$$P_n = \sum_{\substack{k=0, \\ k \neq s}}^{n} c_k f_k(e),$$

where the coefficients c_k are determined from the condition of ensuring the minimum distance between the polynomial under construction and the signal. As shown in [133], in this case

$$c_0 = \frac{\langle f_s(e), f_0(e) \rangle - \sum_{\substack{k=1, \\ k \neq s}}^{n} c_k \langle f_k(e), f_0(e) \rangle}{\langle f_0(e), f_0(e) \rangle},$$

and other coefficients c_k are defined as a solution to the system of linear equations:

$$\sum_{\substack{k=1, \\ k \neq s}}^{n} c_k F_{i,k} = F_{i,s}, \ i = 1, \ldots, n, \ \ i \neq s,$$

where the centered correlates $F_{i,k}$ are also calculated using the appropriate transformations:

$$F_{i,k} = \langle f_i(e), f_k(e) \rangle - \frac{\langle f_i(e), f_0(e) \rangle \cdot \langle f_k(e), f_0(e) \rangle}{\langle f_0(e), f_0(e) \rangle}.$$

A numerical characteristic that can be used in the quality criteria for matching a signal with a selected template, i.e. as a measure of the approximation of the Kunchenko polynomial P_n to the signal $f_s(e)$, can be considered the efficiency coefficient d_n:

$$d_n = \frac{\sum_{\substack{k=1, \\ k \neq s}}^{n} c_k \langle f_k(e), f_s(e) \rangle}{\langle f_s(e), f_s(e) \rangle}.$$

The considered method of recognizing certain patterns by constructing a space with a generating element and searching for the coefficients of the corresponding Kunchenko polynomial can be used in any problem area in which certain characteristic patterns can be identified a priori in the time series.

Thus, having built typical models of the behavior of the series and the intensity of thematic publications during information operations and comparing the patterns obtained on their basis, it is possible to use the method based on Kunchenko polynomials to determine (and prevent) a possible information attack. The work [131] also substantiated the prospects of using this approach in the study of models in the field of statistics and sociology.

Naturally, in practice, focusing only on a single type of sources and mathematical models can lead to a lack of information necessary for decision-making, inaccuracies, and sometimes misinformation. Only the use of complex systems based on the use of numerous sources, databases, mathematical models, along with the above capabilities of content monitoring systems, can guarantee effective information support when countering information operations.

CONCLUSION

The concept of the survivability of an information system implies its ability to perform its functions (informing, influencing, influencing) in a timely manner under the influence of destabilizing factors. In the case of information systems, such factors can be the elimination of individual elements from the information space, the loss of their relevance, the loss of documents. Attracting the attention of the audience to other topics, the generation and development of new information stories can also reduce the relevance of the current information impact. At the same time, from a practical point of view, the origin of the destructive information process plays a much smaller role than its consequences.

Information systems can be both targeted and non-targeted. It is targeted systems that can be generated in the course of active advertising campaigns as distractions, informing and other elements of information operations [3]. At the same time, survivability, which manifests itself as the ability of targeted information systems to perform their main information functions over a given time interval without failures, determines the stability threshold beyond which, without restoring components and functions, an information system can lose its relevance and the possibility of information impact. Because of this and many other factors, the persistence of information systems is essential to information security.

GLOSSARY

Adaptability is the property of a system to adapt to changing conditions of the internal and external environment by using various adaptation mechanisms.

Assortativeness - in the theory of complex networks, a term denoting the predominant connection of network nodes with a high degree of connectivity with each other (the so-called "club of the rich").

Attractor (*from English. attract*) - a compact subset of the phase space of a dynamical system, all trajectories from some neighborhood of which tend to it with time tending to infinity. So, the simplest variants of the attractor are an attractive fixed point (for example, in the problem of a pendulum with friction against air) and a periodic trajectory (example - self-excited oscillations in a circuit with positive feedback).

Wavelets are mathematical functions that allow you to analyze different frequency components of data. Wavelets are local in time and frequency, all functions of one family of wavelets are derived from one using its shifts and stretches along the time axis. All wavelet transforms of any function consider it in terms of oscillations localized in time and frequency.

Deterministic chaos is a phenomenon in which the behavior of a non-linear system appears to be random despite being governed by deterministic laws. The reason for the emergence of deterministic chaos is the instability of the system with respect to the initial conditions and parameters: a change in the initial condition leads to significant changes in the dynamics of the system.

Vitality is the ability of a system to adapt to new, changed and, as a rule, unforeseen situations, to withstand harmful influences, while fulfilling its target function, due to a corresponding change in the structure and behavior of the system. The property of survivability allows a complex system to remain intact in extreme conditions for it, to adapt to them, changing behavior, structure, and often the purpose of functioning. Depending on the class of systems, their complexity, the degree of organization -

lowness, as well as on the chosen level of analysis, the property of viability can be assessed as stability, reliability, adaptability, fault tolerance.

Simulation modeling is a research method in which the system under study is replaced by a model that describes the real system with sufficient accuracy. Experiments are carried out with this model in order to obtain information about the real system. Simulation modeling is a special case of mathematical modeling. There is a class of objects for which, for various reasons, analytical models have not been developed, or methods for solving relative to the obtained model have not been developed. In this case, the mathematical model is replaced by an imitator or simulation model - a logical-mathematical description of the object.

Individual-centric modeling - one of the methods of computer modeling of complex systems - populations. The individual-oriented model describes the mechanisms of interaction of individuals (software agents) with the environment and intrapopulation interactions at the level of individual software agents.

Information impact is the impact on the mass consciousness similar to how the psychological impact affects the individual consciousness. Information impact as a control process is the excitation (inhibition) in the controlled system of such processes that stimulate the choice desired for the influencing party. This method of influencing the enemy does not imply direct disabling of part of the elements of his system, but is the transfer to the enemy of such information that will prompt him to choose the solution necessary for the influencing party.

Information war - a set of measures for informational influence on the mass consciousness to change people's behavior and impose on them goals that are not among their interests, as well as protection from such influences. Purposeful actions taken to achieve information superiority by damaging information, information processes and information systems of the enemy while protecting one's own information, information processes and information systems [133].

Information operations (*Info Ops*, *IO*) - information impact on the mass consciousness (both hostile and friendly), the impact on the information available to the enemy and necessary for him to make decisions, as well as on information and analytical systems of the enemy, including actions aimed at the physical defeat of information and analytical systems, disabling the means of computer and telecommunications infrastructure.

Information space - 1) a set of information resources, technologies for their maintenance and use, information and telecommunication systems that form an information infrastructure; 2) a set of related elements (documents) that form information systems - clusters of related documents.

Information systems - in the context of this work - documentary or content systems - a set of content-related elements of the information space connected to the network. As special cases of information systems, one can consider, for example, thematic collections of documents, websites devoted to a certain issue, or information clusters (plots) - arrays of information messages published on various websites dedicated to one topic or one event.

Information management is the process of developing and applying control actions that are implicit, indirect, informational in nature. The control object is provided with a certain information picture, guided by which, he chooses the line of his behavior, as it seems to him, on his own, i.e. information management is a method of influence that encourages people to behave in an orderly manner, to perform the actions required of them.

An artificial society is a computer model of a society consisting of agents (software analogues of individuals) acting according to certain rules, including interacting with each other.

A cellular automaton is a set of cells that form a certain periodic lattice with given transition rules that determine which determine the state of the cell at the next moment of time through the state of the cells located at a distance no more than a certain distance from it, at the current moment of time. As a rule, automata are considered, where the state is determined by the cell itself and its nearest neighbors.

Content - the content of information resources; any content content of information resources (for example, websites) - texts, graphics, multimedia.

Content analysis is a method of obtaining conclusions by analyzing the content of textual information. Most often it is implemented as a systematic numerical processing, evaluation and interpretation of the form and content of an information source.

Content monitoring is a systematic, continuous scanning and content analysis of information resources.

The clustering coefficient is a value that corresponds to the level of connectivity of nodes in the network. This value shows how many nearest neighbors of a given node are nearest neighbors to each other and is equal to the ratio of the real number of edges that connect the nearest neighbors of this node to the maximum possible.

Latent semantic analysis (*LSA*) is a theory and method for extracting context-dependent meanings of words using statistical processing of large sets of text data. Latent semantic analysis is based on the idea that the totality of all contexts in which a term occurs and does not occur sets a plurality of restrictions, which to a large extent make it possible to determine the similarity of the semantic meanings of terms among themselves. As initial information, LSA uses the terms/documents matrix, which contains the weights of terms in documents.

The logistic equation is an equation that originally appeared when considering the population growth model. The initial assumptions for deriving the equation when considering population dynamics are as follows: the population reproduction rate is proportional to its current size, other things being equal; population reproduction rate is proportional to the amount of available resources, all other conditions being equal
lovia. The second term of the equation reflects competition for resources, which limits population growth. The exact solution of the logistic equation is the logistic function, *the S* -curve.

Mathematical modeling is the process of building and studying mathematical models - mathematical representations of reality.

Multi-agent simulation (*Agent-based Modeling*) - modeling based on (based on) the use of agents. Computer models in which agents are atomic elements.

Modeling - the study of objects of knowledge on their models; building and studying models of real-life objects, processes or phenomena in order to obtain explanations of these phenomena, as well as to predict phenomena that are of interest to the researcher. Modeling can be viewed as the process of creating, applying, using a model. The main functions of the model are the simplification of obtaining information about the properties of an object, the transfer of information and knowledge, the management of objects and processes, their optimization, forecasting, and diagnostics.

Reliability is a complex property of the system, which consists in its ability to perform (under certain operating conditions) the specified functions, while maintaining its main characteristics within certain limits. The most common of the reliability indicators, which are usually probabilistic in nature, are the probabilities of failure-free operation, mean time between failures, availability, etc.

Feedback - the impact of the results of the functioning of the system on the nature of this functioning.

Fault tolerance is the property of a system to remain operational in the event of failure of one or more components.

The behavior of the system is the action of the system in time.

Noise immunity is the property of systems to resist the action of interference (disturbances).

Natural Computing - a direction within the framework of the concept of multi-agent modeling, combining mathematical methods that contain decision-making principles similar to mechanisms implemented in nature. For example, mimicking the self-organization of an ant colony (or termite colonies) forms the basis of the so-called ant colony optimization algorithms, one of the promising methods of natural computing.

Psychological impact - impact on the psyche of individuals, which can be carried out by various means: information, military, economic, political. Psychological impact is divided into the following types: formational - psychological, psychogenic, psychoanalytic, neurolinguistic, psychotronic, psychotropic.

Reductionism (*from lat. - Reductio*) - a methodological principle according to which complex phenomena can be fully explained using the laws inherent in simpler phenomena (for example, sociological phenomena are explained by biological or economic laws).

Synergetics (*from the Greek συν* - jointly *and* εργος - acting) is an interdisciplinary area of scientific research, the task of which is to study natural phenomena and processes based on the principles of self-organization of systems (consisting of subsystems) ; "... a science that studies the processes of self-organization and the emergence, maintenance, stability and decay of structures of the most diverse nature...".

System (*from other - Greek* σύστημα - *combination*) - a set of interconnected elements, isolated from the environment and interacting with it as a whole.

A system of cellular automata is a set of mathematical objects that represent a homogeneous grid, each cell of which (cellular automaton) can be in one of the possible states. Cell states are synchronously updated at each simulation step in accordance with certain transition rules; in the general case, there can be an infinite number of such rules , which corresponds to the number of subsets of a countable set.

The system effect is the irreducibility of a system to the sum of the properties of its components.

Scaling (*from the English. Scaling)* - scale invariance, self-similarity. This property is used, in particular, to represent a function of two variables as a function of one.

A scalogram is a map of transformation coefficients.

Skeleton are the lines of local extrema of the scalogram surface.

A complex system is a system consisting of many interacting components (subsystems), as a result of which a complex system acquires new properties that are absent at the subsystem level and cannot be reduced to the properties of the subsystem level.

Complex network - a network (graph) with non-trivial topological features that are not characteristic of simple networks, such as lattices or random graphs. The study of complex networks is a field of scientific research that has emerged from the study of real world networks such as computer and social networks.

Structure (*from lat. - structura - structure*) - a relatively stable fixation of connections between the elements of the system. In natural science, structure is the internal structure of something, hidden by the external form of the object. The structure is associated with the orderliness of relations that connect the elements of the system. Structures are simple, complex, hierarchical.

Text corpus - an array of texts collected in accordance with certain principles, marked up according to a certain standard and provided with a specialized search engine. In some cases, the text corpus of the first order is called arbitrary arrays, united by some common feature. The development, creation and use of text (linguistic) corpora is handled by a special section of linguistics - corpus linguistics.

Thematic information flow is a sequence of messages corresponding to a specific topical query.

Term - a word or a constant phrase. In mathematical logic, the concept of "term" is widely used as a "symbolic expression".

The bifurcation point is a critical state of the system, at which the system becomes unstable with respect to fluctuations and uncertainty arises: whether the state of the system will become chaotic or whether it will move to a new, more differentiated and higher level of order.

The stability of the system is the ability of the system to return to its initial state after the end of the influence that brought the system out of this state. The active preservation by a system of certain characteristics, regardless of whether they play any role in the overall system.

Vulnerability is a parameter that characterizes the possibility of causing damage to the described system of any nature by one or another external means or factors.

The phase plane is a coordinate plane in which any two variables (phase coordinates) are plotted along the coordinate axes, which uniquely determine the state of the second-order system. The phase plane is a special case of the phase space, which can have a large dimension. In the physics of oscillations, the value of the parameter *x is plotted on the abscissa of the phase plane* , and the first derivative *of x* with respect to time is plotted on the ordinate.

Phase space is the set of all states of the system at a fixed point in time. Each possible state of the system corresponds to a point in the phase space. The essence of the concept of phase space lies in the fact that the state of an arbitrarily complex system is represented in it by one single point, and the evolution of this system is represented by the displacement of this point.

Fractal (*from lat. Fractus - crushable, consisting of fragments*) - infinitely itself about a similar (exactly or approximately) object (plural), each part of which is repeated when zooming out. Another definition is a geometric object whose Hausdorff–Besikovich dimension is non-integer. The following definition is also possible: a fractal is a self-similar set of non-integer dimension.

Fractal analysis is a modeling method using the theory of fractals, which consists in studying the fractal dimension and other fractal properties of signals, data sets, objects.

Purposeful behavior of the system - the desire of the system to achieve the goal of functioning.

Integrity is the relative independence of the system from the environment and other similar systems.

Information integrity is a term in computer science and telecommunications theory that means that the data is complete; the condition that the data has not been modified by any operation on it, be it transmission, storage, or presentation.

The goal of the system is a state that is preferable for the system, i.e. some final state to which the system tends due to its structural organization.

The evolution of the system is the change in the structure of the system over time.

Element (*from Lat. - elementum - element*) - a further indecomposable (in a given system with a given method of consideration and analysis) component.

Emergence in systems theory - the presence of any system of special properties that are not inherent in its subsystems and blocks, as well as the sum of elements that are not connected by special system-forming links; irreducibility of system properties to the sum of properties of its components; synonym - "systemic effect".

BIBLIOGRAPHY

1. *Shelton C.* A Framework for Scalable Analysis and Design of System-Wide Graceful Degradation in Distributed Embedded Systems / C. Shelton, P. Koopman, W. Nace // Eighth IEEE Int. Workshop on Object-Oriented Real-Time Dependable Systems (WORDS 2003). — Guadelajara (Mexico), Jan. 2003. - 8 p.
2. *Dodonov A.G.* Introduction to the theory of survivability of computing systems / A.G. Dodonov, M.G. Kuznetsova, E.S. Gorbachik. - K.: Nauk. d umka, 1990. - 184 p.
3. *Gorbulin V.P.* Informational operations and security of support: threats, protides, modeling: monograph / V.P. Gorbulin, O.G. Dodonov, D.V. Lande. — K.: Intertekhnologiya, 2009. — 164 p.
4. *Broder A., Kumar R., Maghoul F. et al.* Graph structure in the Web / A. Broder, R. Kumar, F. Maghoul et al. //Proc. 9th I nt. World Wide Web Conf. on Computer Networks: The International Journal of Computer and Telecommunications Networking. - Amsterdam, 2000. - P. 309–320.
5. *Dodonov A.G.* Vitality of information plots as dynamic documentary systems / A.G. Dodonov, D.V. Lande // Registration, collection and processing. danich. - 2010. - V. 13, No. 2. - S. 88–102.
6. *Dodonov O.G.* Survivability of folding systems: analysis and modeling: Navch. settlement ib. at 2 hours / O.G. Dodonov, M.G. Kuznetsova, O.S. Gorbachik. - K.: NTUU "KPI", 2009. - 264 p.
7. Modeling the revolution in military affairs, autumn // JFQ. — Winter 1998–99.
8. *Ivanov S.A.* The World System of Scientific Communication as an Information Space // Libraries and Associations in a Changing World: New Technologies and New Forms of Cooperation: 8th Intern. conf. "Crimea 2002": Proceedings of the Conf., Sudak, June 9–17, 2001 г. - M., 2001. - T. 1. - S. 1123–1126.
9. *Langton C.* Artificial Life: ser. SFI Studies in the Sciences of Complexity. - Red-Wood City, CA: Addison-Wesley, 1989. - P. 1–47.
10. *Epstein JM* Remarks on the foundations of agent-based generative social science // SFI (Santa Fe Institute) Working Paper. DOI: SFI-WP 05-06-024, SFI Working Papers, 2005.
11. *Karpov Yu.G.* Simulation modeling of systems. Introduction to modeling with AnyLogic 5. - BHV-Petersburg, 2006. - 400 p.

12. *Karpov Yu.G.* The study of modern paradigms of simulation modeling in the AnyLogic environment // Computer tools in education. - 2005. - No. 4. - S. 3-14.
13. *Borshchev A.V.* Practical agent-based modeling and its place in the analyst's arsenal // Exponenta Pro. - 2004. - No. 3-4.
14. *Shtovba S.D.* Ant algorithms // Exponenta Pro. Mathematics in applications. - 2003. - No. 4. - S. 70–75.
15. *Axelrod R.* Modeling Security Issues of Central Asia // R. Gerald Ford SchoolofPublic Policy Universityof Michigan(Project on "Security in Central Asia". - June 2004. - US Govt. Contract # 2003 * H513400 * 000).
16. *Lustick I.*PS-I: A user-friendly agent-based modeling platform for testing theories of political identity and political stability // J. Artificial Societies and Social Simulations. - 2002. - Iss. 5, N 3.
17. *Forrester JW* System dynamics and the lessons of 35 years // A Systems–Based approach to Policymaking / Ed. by UB De Green UB - Boston: Kluwer, 1995. - P. 199 - 239.
18. *Grimm V.* Ten years of individual-based modeling in ecology: what have we learned and what could we learn in the future // Ecological Modeling. - 2002. - **115** , N 2 - 3. - P. 129-148. _
19. *Epstein JM, Axtell R.* Artificial societies and generative social science // Artificial Life and Robotics. - 1997. - N 1. - P. 33 - 34.
20. *Neumann J.* Theory of self-reproducing automata. — M.: Mir, 1971. — 382 p.
21. *Theory* and applications of cellular automats / Ed. ByS.A. Wolfram. - Singapore: World Scientific, 1986.
22. *Lande D.V.* Modeling of electoral processes based on the concept of cellular automata / D.V. Lande, V.N. Furashev // Open Information and Computer Technologies. Kharkov: NAKU "KhAI", 2007. - Issue. 36 - S. 17-34.
23. *Plotinsky Yu.M.* Models of social processes. - Ed. 2nd. — M.: Logos, 2001. — 296 p.
24. *Plotinsky Yu.M.* Mathematical modeling of the dynamics of social processes. - M.: Publishing house of Moscow State University, 1992. - 133 p.
25. *Gardner M.* Mathematical leisure. — M.: Mir, 2000. — 448 p.
26. *Toffoli T., Margolus N.* Machines of cellular automata. — M.: Mir, 1991. — 280 p.

27. *Brown TA* Nonlinear politics // Chaos Theory in the Social Sciences / Eds. LD Kiel, E. Elliot. Ann Arbor. - The Univ. of Michigan-Press, 1996. - P. 119-137.

28. *Casti JL* Searching for Certainty. New York: W.Morrow, 1990.

29. *Schelling TS* Models of segregation. American Economic Review // Pap. and Proc. - 1969. - **59.** -p. 488 - 493.

30. *Bhargava S.C., Kumar A., Mukherjee A.* A stochastic cellular automata model of innovation diffusion // Technol. Forecast. and So c. change. - 1993. - **44** , No. 1. — P. 87–97.

31. *Tiebout Ch.M.* A Pure theory of local expenditures // J. Political Economy. —1956. - **64** , N 5. -P. 416-424. _

32. *Lande D.V.* Information diffusion model // Information technologies and security. Information security management: Sat. scientific tr. Institute of Problems of Registration of Information of the National Academy of Sciences of Ukraine. - 2007. - Issue. 10. - S. 51-67.

33. *Ryabinin I.A., Cherkesov G.N.* Logical and probabilistic methods for studying the reliability of structurally complex systems - M.: Radio and communication, 1984. - 238 p.

34. *Reinshke K.* Models of reliability and sensitivity of systems. — M.: Mir, 1979. — 454 p.

35. *Reinshke K., Ushakov I.A.* Estimation of system reliability using graphs. - M.: Radio and communication, 1988. - 208 p.

36. *Dodonov A.G., Fleitman D.V.* On the issue of information systems security // Information Technologies and Security: Zb. Sciences. p rats. - K., 2004. - VIP. 6. - S. 26–29.

37. *Averyanov V.T.* Selection and study of the properties of the indicator of the level of survivability of the fire extinguishing system // Problems of risk management in the technosphere. - 2009. - V. 9-10. — No. 1–2. — pp. 55–62.

38. *Ashby W.R.* Principles of self-organization. — M.: Mir, 1966. — 622 p.

39. *Petrushenko L.A.* The unity of consistency, organization and self-promotion. - M.: Thought, 1975. - 286 p.

40. *Petrov Yu.P.* Information and entropy in cybernetics. - L.: LGU, 1989. - 84 p.

41. *Shambadal P.* Development and applications of the concept of entropy. — M.: Nauka, 1967. — 280 p.

42. *Bestuzhev-Lada I.V.* Forecasting workbook. - M.: Thought, 1982. - 430 p.

43. *Survivable* network systems: an emerging discipline CMU/SEI-97-TR-013 ESC-TR-97-013. —Software Engineering Institute Carnegie Mellon University Pittsburgh. May 1999.

44. *Newman MEJ* The structure and function of complex networks // SIAMRev. - 2003. - **45.** - P. 167 - 256.

45. *Dorogovtsev SN, Mendes JFF* Evolution of networks: from biological networks to the Internet and WWW. — Oxford UniversityPress, 2003.

46. *Erdős, P., Renyi A.* On Random Graphs. I // Publ. Math. - 1959. - **6.** - P. 290 - 297.

47. *Erdős P., Rényi A.* On the evolution of random graphs // Publ. Math. Inst. hungar. Acad. —1960. — Sci. 5. - P. 17 - 61.

48. *WattsDJ, Strogatz SH* Collective dynamics of "small-world" networks // Nature. - 1998. - **393.** - P. 440 - 442.

49. *Albert R., Jeong H., Barabasi A.* Attack and error tolerance of complex networks // Nature. - 2000. - **406.** - P. 378 - 382.

50. *Gol'dshtein V., Koganov GA, Surdutovich GI* Vulnerability and hierarchy of complex networks // Phys. Rev. Lett. — 2004.

51. *Zhou S., Mondragon RJ* The rich-club phenomenon in the internet topology // Commun. Lett. IEEE. - 2004. - **8.** - P. 180-182.

52. *Boguna M., Pastor-Satorras R., Vespignani A.* Statistical mechanics of complex networks / Lecture and notes in physics. — Springer Berlin. - 2003. - P. 127 - 147.

53. *Newman MEJ* Assortative mixing in networks, Phys. Rev. Lett. - 2002. - **89** (208701).

54. *Barabási A., Albert R.* Emergence of scaling in random networks // Science. - 1997. - **286.** - P. 509 - 512.

55. *Bjorneborn L., Ingwersen P.* Toward a basic framework for webometrics // J. Amer. soc. inform. sci. and Technol. - 2004. - **55** (14). - P. 1216-1227.

56. *Milgram S.* The small world problem // Psychology Today. — 1967. - **2.** - P. 60 - 67.

57. *WattsDJ, Strogatz SH* Collective dynamics of " small-world " networks // Nature. - 1998. - **393.** - P. 440-442.

58. *Newman MEJ, WattsDJ* Scaling and percolation in the small-world network model, Phys. Rev. E. - 1999. - **7332.**

59. *Snarsky A.A.* Transport processes in macroscopic disordered media: From mean field theory to percolation. / A.A. Snarsky, I.V. Bezsudnov, V.A. Sevryukov. - M.: URSS, Publishing house LKI, 2007. - 304 p.

60. *Tarasevich Yu.Yu.* Percolation: theory, applications, algorithms. - M., URSS, 2002. - 112 p.

61. *Broadbent SR, Hammersley JM* Percolation processes: I. Crystals and MAZES // Proc. Cambridge Phil. soc. - 1957. - P. 629 - 641.

62. *Cohen R., Erez K., ben-Avraham D., Havlin S.* Resilience of the Internet to Random Breakdown // Phys. Rev. Lett. - 2000. - **85.** — P. 4626.

63. *Snarskii A.* FreeBSD Stack Integrity Patch. 1997 URL: ftp://ftp.lucky.net/pub/unix/local/libc-letter

64. *The* Deep Web: Surfacing Hidden Value, 2000 BrightPlanet.com LLC, 35 p. URL: http://www.dad.be/library/pdf/BrightPlanet.pdf

65. *Braichevsky S.M., Lande D.V.* Modern information flows: topical issues // Scientific and technical information. Ser. 1. - 2005. - Issue. 11. - S. 21-33.

66. *Dodonov O.G.* Information flows in global computer networks / O.G. Dodonov, D.V. Lande, V.G. Putyatin. - K.: Nauk. Dumka , 2009. - 295 p.

67. *Lande D.V., Furashev V.M., Grigor'ev O.M.* The software and hardware complex of informational and idtrimki adopted a solution: Scientific and methodological guide. - K.: Engineering, 2006. - 48 p.

68. *Lande D.V.* The structure of the news Web-space // Scientific and technical information. Ser. 2. - 2006. - No. 8. - S. 17–20.

69. *Lande D.V.* Fundamentals of Information Flow Integration: Monograph. - K.: Engineering, 2006. - 240 p.

70. *Grigoriev A.N., Lande D.V.* Adaptive interface for refining queries to the InfoStream content monitoring system // Computer Linguistics and Intelligent Technologies: Proceedings of the Intern. conf. "Dialogue-2005" (Zvenigorod, June 1–6, 2005 г.). — M.: Nauka, 2005. — S. 109–111.

71. *Luisi PL* Emergence in Chemistry: Chemistry as the Embodiment of Emergence // Found. Chem. - 2002. - **1.** - P. 183-200.

72. *Feigenbaum MJ* Quantitative universality for a class of nonlinear transformations // J. Stat. Phys. - 1978. - **19.** — P. 25–52.

73. *Potev M.I.* Electronic textbook on discipline: "Concepts of modern natural science". — St. Petersburg State University ITMO, URL: http://de.ifmo.ru/bk_netra/start.php?bn=12

74. *Prigogine I., Stengers I.* Order out of chaos. - M., 1986. - 432 p.

75. *Arnold V.A.* Theory of catastrophes. — M.: Nauka, 1990. — 128 p.

76. *Kokovin S.G.* Lectures on game theory and political science. Part 1. Introduction to game theory. — 66 p. (http://www.math.nsc.ru/~mathecon/ Kokovin/m1tigran.pdf)

77. *Rastorguev S.P.* The philosophy of information warfare. - M.: Vuzovskaya book, 2001. - 468 p.

78. *Intrilligator M.* Mathematical optimization methods and economic theory. — M.: Progress. - 1975. - 607 p.

79. *Stackelberg H. von.* Marktform and Gleichgewicht. Faksimile der 1934 inWien und Berlinerschienenen Erstausgabe. Reihe: Klassiker der Nationalökonomie VI, 142 S., Gebunden.

80. *Fursova P.V., Levich A.P., Alekseev V.L.* Extremal principles in mathematical biology // Successes of modern biology. - 2003. - T. 123, No. 2. - S. 115-117.

81. *Webb JN* Hamilton's variational principle and ecological models // Ecol. Modelling. - 1995. - 80. - P. 35.

82. *Pritz A.K.* The principle of stationary states of open systems and population dynamics. - Kaliningrad, 1974. - 123 p.

83. *Lurie D., Valls J., Wagensberg J.* Thermodynamic approach to biomass distribution in ecological systems // Bull. Math. Biol. 1983. - **45.** — P. 869.

84. *Levich A.P.* The structure of ecological communities. - M.: Publishing House of Moscow. un-ta, 1980. - 181 p.

85. *Svirezhev Yu.M.* Phenomenological thermodynamics of interacting populations // Zhurn. total _ b biology. - 1991. - T. 52, No. 6. - S. 840.

86. *Khanin M.A.* Energy and criteria for the optimality of ontogenetic processes. Mathematical biology of development. — M.: Nauka, 1982. — 177 p.

87. *Khanin M.A., Dorfman N.L., Kukharov I.B., Levadny V.G.* Extreme principles in biology and physiology. — M.: Nauka, 1978. — 256 p.

88. *Evdokimov E.V.* Problems of Regular Behavior and Deterministic Chaos in the Basic Models of Population Dynamics (Theory and Experiment) // Abstract of the thesis. dis. for an apprenticeship doctorate in biological sciences. - Krasnoyarsk, 1999. - 40 p.

89. *Malkov S.Yu.* Mathematical modeling of historical dynamics: approaches and models. — M.: RGSU, 2004. — 188 p.

90. *Zhabotinsky A.M.* concentration fluctuations. — M.: Nauka, 1974. — 179 p.

91. *Voltaire W.* Mathematical theory of the struggle for existence. — M.: Nauka, 1976. — 288 p.

92. *Gause G.F.* Struggle for existence. - M: URSS, 2002. - 160 p.

93. *Gause G.F.* Ecology and some problems of the origin of species // Ecology and evolutionary theory. - L., Nauka, 1984. - S. 5–108.

94. *Chernavsky D.S., Chernavskaya N.M., Malkov S.Yu., Malkov A.S.* Mathematical modeling of geopolitical processes // Strategic stability. - 2002. - No. 1. - S. 60–66.

95. *Lande D.V.* Search for knowledge on the Internet. Professional work. - M.: Dialectics, 2005. - 272 p.

96. *Furashev V.M., Lande D.V.* Informational Operations of the Crisis of the Prism of the Monitoring System and Integration of Internet Resources // Legal informatics. - 2009. - No. 2 (22). — pp. 49–57.

97. *Hartley R.* Transmission of information. Theory of information and its applications. — M.: Fizmatgiz, 1959.

98. *Chui K.* Introduction to wavelets. — M.: Mir, 2001. — 416 p.

99. *Astafieva N.M.* Wavelet analysis: fundamentals of theory and examples of application // Uspekhi fizicheskikh nauk. - 1996. - T. 166, No. 11. - S. 1145-1170.

100. *Davydov A.A.* Wavelet analysis of social processes // Sociological research. - 2003. - No. 11. - S. 97–103.

101. *Davydov A. A.* System sociology. - M.: KomKniga, 2006. - 192 p.

102. *Mandelbrot B.* Fractal geometry of nature. - M.: Institute of Computer Research, 2002. - 656 p.

103. *Feder E.* Fractals. — M.: Mir, 1991. — 254 p.

104. *Peng CK, Buldyrev SV, Havlin S., Simons M., Stanley HE, Goldberger AL* Mosaic organization of DNA nucleotides. // Phys Rev E. - 1994. - **49** (2). - P. 1685-1689.

105. *Lande D.V., Snarsky A.A.* Dynamics of the deviation of elements of a series of measurements from local linear approximations // Registry, collection and processing. danich. - 2009. - V. 11, No. 1. - S. 27–32.

106. *Mandelbrot BB* Fractals and m ultifractals // Noise, Turbulence and Galaxies, Selecta. - New York: Springer-Verlag , 1989. - **1.**

107. *Pavlov A.N., Sosnovtseva O.V., Ziganshin A.R.* Multi-fractal analysis of the chaotic dynamics of interacting systems // Izv. universities, Applied Nonlinear Dynamics. - 2003. - V. 11, No. 2. - S. 39–54.

108. *Snarsky A.A., Lande D.V., Braichevsky S.M., Darmokhval A.T.* Relevance Distribution Properties in Documentary Arrays // Proceedings of the 9th All-Russian Scientific Conf. "Digital Libraries: Advanced Methods and Technologies, Digital Collections" - RCDL'2007, Pereslavl-Zalessky, Russia, 2007. - P. 151–155.

109. *Information* operations roadmap. — DoD US, Oct. 30, 2003. - 78 p.

110. *Hill JMD, Surdu JR, Ragsdale DJ, Schafer, JH* Anticipatory planning in information operations // Systems, Man, and Cybernetics: 2000 IEEE Int. Conf. - 2000. - Vol. 4. - P. 2350-2355.

111. *Burke MM* Knowledge Operations: above and beyond Information Operations // 6th Int. Command and Control Research and Technology. June 19–21, 2001

112. *Lasswell HD* The structure and function of communication in society // The Communication of Ideas / Ed. by L. Bryson. - New York: Harper and Brothers, 1948.

113. *Schramm W., Roberts DF* The Process and Effects of Mass Communication. - Univ. of Illinoise Press, 1974.

114. *Osgood Ch.E.* Psycholinguistics. A survey of theory and research problems // Supplement to the Int. J.Amer. Linguistics. - 1954. - **20** , N 4. - Baltimore: Waverly Press, 1954.

115. *Latane B.* The psychology of social impact // American Psychologist. - 1981. - **33.** - P. 343 - 356.

116. *Latane B., Nowak A.* Causes of polarization and clustering in social groups // Progr. commun. sci. - 1997. - **13.** - P. 43 - 75.

117. *Nowak A., Szamrej J., Latané B.* From private attitude to public opinion: A dynamic theory of social impact // Psychological Review. - 1990. - **97.** — R. _ 367 - 376.

118. *Lewenstein M., Nowak A., Latane B.* Statistical mechanics of social impact // Phys. Rev. _ A. — 1993. — **45.** — R. _ 763 - 776.

119. *Kacperski K., Holyst JA* Phase transitions as a persistent feature of groups with leaders in models of opinion formation // Physica A. Statistical Mechanics and its Applications. - 2000. - **287** , Iss. 3 - 4. - P. 631-643.

120. *Sobkowicz P.* Effect of leader's strategy on opinion formation in networked societies. — 2003. - URL http://arxiv.org/pdf/condmat/ 0311566.

121. *Sobkowicz P.* Opinion formation in networked societies with strong leaders. — 2003. - URL http://arxiv.org/pdf/condmat/0311521.

122. *Kulba V.V., Malyugin V.D., Shubin A.N., Vus M.A.* Introduction to information management. Educational and methodical edition. - St. Petersburg: Publishing House of St. Petersburg University. 1999. - 116 p.

123. *Kononov D.A., Kulba V.V., Shubin A.N.* Basic concepts of modeling information management in social systems // Tr. intl. Scientific and Practical Conf. "Theory of Active Systems". - M.: Inst. problems of management. V.A. Trapeznikova RAN, 2003. - T. 2. - S. 125–129.

124. *Kulba V.V., Kononov D.A., Kosyachenko S.A., Shubin A.N.* Methods of formation of scenarios for the development of socio-economic systems. - M.: SINTEG, 2004. — 296 p.

125. *Chkhartishvili A.G.* Game-theoretic models of information management. - M.: CJSC "PMSOFT", 2004. - 227 p.

126. *Bak P.* How nature works: The Science of Self-Organized Criticality. - New York: Springer-Verlag Inc., 1996.

127. *Kapitsa S.P., Kurdyumov S.P., Malinetsky G.G.* Synergetics and forecasts of the future. — M.: Nauka, 1997. — 288 p.

128. *Van Raan AFJ* Fractal geometry of information space as represented by cocitation clustering // Scientometrics. - 1991. - **20** , N 3. - P. 439 - 449.

129. *Ivanov S.A.* Stochastic fractals in Informatics // Scientific and technical information. Ser. 2. - 2002. - No. 8. - S. 7-18.

130. *Lande D.V.* Fractal properties of thematic information flows from the Internet // Registration, collection and processing. danich. - 2006. - V. 8, No. 2. - S. 93–99.

131. *Chertov O.R.* Polynomial of Kunchenko for recognition of images // VISNIK NTUU KPI. Ser. Informatics, management and calculation technology. - 2009. - No. 51. - S. 103–108.

132. *Kunchenko Yu.P.* Approximations in space with a generating element. - K.: Nauk. Dumka, 2003. - 243 p.

133. *Pocheptsov G.G.* information wars. M.: Refl-book, K.: Vakler, 2000. - 576 p.