ISSN 2782-2427

CONTROL SCIENCES

5/2023

[1]Russian Academy of Sciences.

CONTROL SCIENCES
Scientific Technical Journal

6 issues per year
ISSN 2782-2427
Open access

Published since 2021

Original Russian Edition
Problemy Upravleniya
Published since 2003

FOUNDER AND PUBLISHER
V.A. Trapeznikov
Institute of Control Sciences
of Russian Academy of Sciences

Editorial address
65 Profsoyuznaya st., office 410,
Moscow 117997, Russia

☎/🖹 +7(495) 198-17-20, ext. 1410

✉ pu@ipu.ru

URL: http://controlsciences.org

Published: November 15, 2023

Registration certificate of
Эл № ФС 77-80482
of 17 February 2021
issued by the Federal Service
for Supervision of Communications,
Information Technology, and Mass
Media

CONTROL SCIENCES

5.2023

CONTENTS

DOI: http://doi.org/10.25728/cs.2023.5.1

DISTRIBUTED INTELLIGENCE OF MULTI-AGENT SYSTEMS. PART I: Basic Features and Simple Forms

Yu.L. Slovokhotov[1,2] and D.A. Novikov[1]

[1]Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
[2]Department of Materials Science, Moscow State University, Russia

✉ yurislovo@yandex.ru, ✉ novikov@ipu.ru

Abstract. This multi-part survey is devoted to the empirical characteristics and manifestations of intelligence as the capability of an autonomous agent to perceive, process, and use information in order to achieve its goal. Part I of the survey briefly describes the most significant attributes of "proto-intelligent" and intelligent systems, together with the main features of distributed intelligence (DI), individual human intelligence (HI), and artificial intelligence (AI). An inseparable connection is emphasized between the DI of human organizational and social systems and individual human intelligence. The simplest forms of swarm intelligence are considered as examples, and the key factors determining the effectiveness of DI in such multi-agent systems are presented, including the structure of interactions between agents, their collective goal-setting, external information recording, compression, and processing, and the standard "images" of external influences. Their combination creates the performance of a multi-agent system far beyond the capabilities of its individual agents. In part II of the survey, different forms of collective intelligence in human social systems will be analyzed and all known types of intelligence will be generally classified.

Keywords: multi-agent systems, distributed intelligence, swarm intelligence, organizational systems, collective intelligence.

INTRODUCTION

The analysis of human intellectual activity and its mathematical and technical modeling are the main directions of cognitive psychology, information technology, cybernetics, robotics, and other sciences [1]. Based on huge data in these fields, the key common features of *human intelligence* (HI) and various computer implementations of *artificial intelligence* (AI) have been found. Adjacent directions include research and modeling of cooperative dynamics in various multi-agent systems (biological, social, economic, and organizational systems) and in groups of autonomous technical devices. This survey considers information processing and use in such systems, the processes generally related to the manifestations of *distributed intelligence* (DI).

Collective information processing and its use by organizational systems (OSs) are well known [1]. Examples are competition of firms, the combat operations of military units, participation of political parties in election campaigns, and other processes. But everyday manifestations of distributed intelligence in multi-agent systems are much broader. They include, among others, the separation of pedestrian traffic into opposite-direction flows (which increases the capacity of sidewalks and tunnels), slowing down of automobile transport on congested streets (where time losses from traffic jams are outweighed by safety considerations [2, 3]), collapsing sales of assets on the stock exchange (which indicates the onset of a crisis), etc. Unlike standard control mechanisms for OSs, the "consensus" of participants in such systems is established in a decentralized way without general discussions: their coordinated behavior is primarily based on mutual connectivity and the structure of inter-agent interactions.

Numerous manifestations of DI without conscious processing of information by individuals[1] are well

[1] Whenever no confusion occurs, "individual" will be preferably replaced by "agent."

known for social insects [4, 5], flocks of birds and schools of fish [6, 7], pedestrian flows [2, 8], and formations of drones [8, 9]. This also includes other forms of human activity such as the market and stock exchange, where the desire of participants to maximize profits is transformed into a collective value of goods [10]. Note that the DI of different multi-agent systems is not reducible to standard schemes of individual or collective decision-making or hierarchical or network control: it represents an independent insufficiently studied aspect of cooperative dynamics [11].

This survey considers the main modern kinds of DI in systems of interconnected *agents*, including biological, technical, social, and organizational ones. Without delving into the vast field of intelligence modeling, we will list the main kinds of DI and consider their analogies with the intelligent behavior of humans and animals[2] as well as with some kinds of AI.

Due to the large volume of material, the survey will include two parts. In part I (here), the general features of all kinds of intelligence are discussed, including examples of their partial realization in "proto-intelligent" automatic control systems and in existing AI models. This part describes in detail the simplest forms of *swarm intelligence* in groups of biological individuals (social insects, schools of fish, and flocks of birds) and in technical systems (swarm intelligence in robotics), including the simulation of DI in agent-based models and in modern computing algorithms. Part II of the survey will deal with different kinds of DI in human communities, the integration of HI and AI into "collective intelligence," and the general classification of all kinds of intelligence.

1. EMPIRICAL CHARACTERISTICS OF INTELLIGENCE

The vast literature on human intellectual activity paradoxically lacks a universally accepted definition of intelligence. For example, in the fundamental monograph [12], all five chapters of the first volume were devoted to different approaches to describing and explaining this phenomenon. The phenomenon of intelligence is studied in a wide range of disciplines from engineering, biology, and psychology to philosophy, with cognitive sciences as the center of gravity. Most definitions of intelligence in all these fields, with varying degrees of detail, describe it empirically as the ability to use limited external information to formulate and solve nonstandard tasks facilitating the individual's adaptation to a changing environment. Undoubtedly, they are based on human intelligence, but a much wider range of objects and systems, including technical, biological, and social ones, meets this characterization; see above.

Figure 1 presents different kinds and types of intelligence and its carriers discussed in the modern literature. Being far from complete, this compact scheme does not show logical connections of similar manifestations of intellectual activity in individuals and communities of different nature (swarm intelligence of lower animals and crowdsourcing in human society, natural computing algorithms, the emergent AI of multi-agent systems, and much more), which are actively discussed in the literature and will be considered below. The only exception is made for the diverse implementations of collective human intelligence in various social structures (the dashed arrow), which constitute the main subject of our discussion. Different formal models are used in describing different kinds of intelligence; we will further analyze their common features. The largest-scale processes in social systems, highlighted in gray on the scheme, are most often discussed at the descriptive level in the humanities [13]. But they certainly manifest many kinds of collective intelligence and are therefore also included in the classification.

While manifestations of DI in weakly structured biological and human communities are of great academic interest (see the books [10, 11] and references therein), "intelligent" technical and organizational multi-agent systems are predominantly addressed by control science [14, 15] and applied engineering disciplines [16, 17]. In the last decade, the term "*emergent intelligence*" (EI) has gained popularity in this field: the "spontaneously arising" and "nascent" intelligence in natural and artificial complex systems. This term often refers to flexible multi-agent planning systems for production and logistics, the results of which obviously exceed the individual capabilities of the units forming them. This term is very frequently used to describe the DI of economic, transportation, and other large systems; see [16] and other books in this series.

In applications, somewhat advertizing phrases ("intelligent terminal," "intelligent system," etc.) are widespread to describe an advanced, versatile, and flexible technical or organizational scheme. This is especially true in the vast field of *robotics* [18], where the term "intelligence" often serves as a metaphor far beyond its interpretations in psychology or cognitive science. Indeed, automatic control systems of autonomous technical devices exhibit some essential features of *all* kinds of intelligence and, in this sense, can be called "proto-intelligent"; see the considerations below.

[2] In the biological literature, it is commonly referred to as the cognitive abilities of animals.

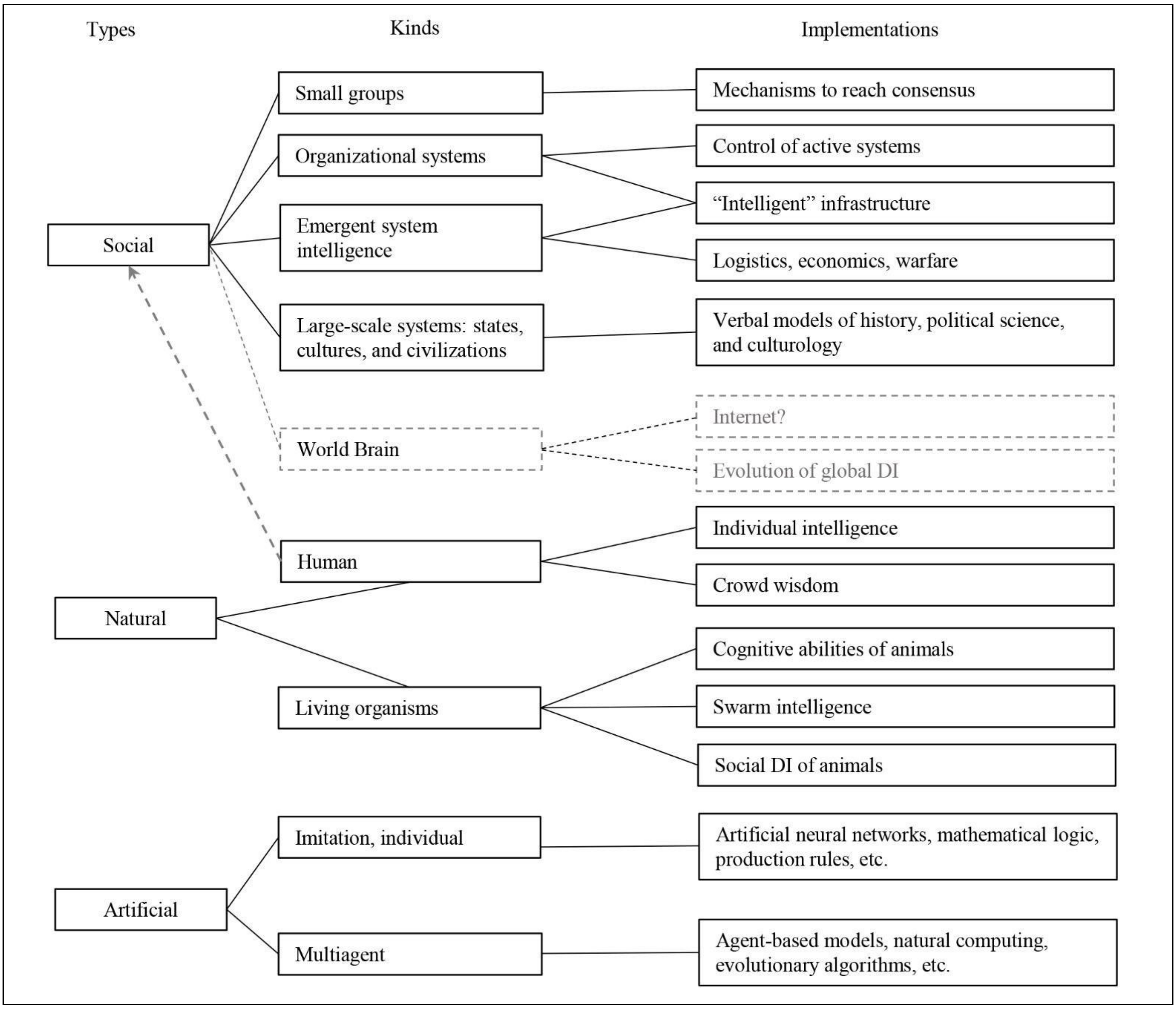

Fig. 1. The types, kinds, and implementations of intelligence.

The characteristic paper [19], the opener of the proceedings of the 2006 American Symposium on *Artificial General Intelligence*[3], provided 18 general scientific and 35 psychological definitions of intelligence used in the literature along with 18 different definitions of AI generally accepted in the same sense. According to the Encyclopædia Britannica [20], also cited in [19], most of the controversy in the field stems from attempts to provide a precise definition of intelligence. Against this background, the trends of the most serious sources to provide trivial characterizations ("intelligence as a measure of an agent's ability to achieve its goals in a widely varying environment" [19], "artificial intelligence: the ability of a digital computer or a computer-controlled robot to perform tasks commonly associated with intelligent beings" [20]) and even somewhat parodic characterizations ("in view of all this complexity, we use a simple definition of collective intelligence: cooperative actions by a group of individuals that appear intelligent" [21]) seem unsurprising. In most of the formulations used, intelligence appears as an intuitive entity, and its selected aspects are the focus of a particular study. Such ambiguities introduce uncertainty into the discussion of intelligence as a phenomenon, simultaneously creating some freedom in the choice of terms to analyze it.

The features of human intelligence frequently mentioned in its most common definitions, albeit not all of them, include the following: cognitive abilities, goal-setting, learnability (including the ability to learn from

[3] We do not discuss Artificial Narrow Intelligence, actually applied in practice, and so far hypothetical Artificial General Intelligence, comparable to human intelligence.

experience), adaptation to new situations, productive use of abstract notions (concepts and images), reflexion, application of knowledge for purposeful impact on the environment, information exchange, and use of speech. Besides the amorphousness of several terms (knowledge, concepts, and speech), also having many different meanings in different contexts, they neglect such kinds of intelligence where agents do not use abstract concepts and speech ("intellectual" abilities of animals), are not inclined to reflexion (social insects), or have no cognitive abilities in the usual sense (formations of robots and "swarm robotics"; see the book [18]). Therefore, it is most reasonable to start with the broadest possible definition of "*proto-intelligence*," satisfied by most objects that are (sometimes metaphorically) called intelligent in the specialized literature. Particular kinds of DI will be discussed by detailing appropriate concepts and narrowing the scope of their applicability.

As a necessary, albeit insufficient, condition for the presence of intelligence, we will call ***the ability of an autonomous agent to tend to a certain state (goal) by perceiving, processing, and using external information to achieve it***. Such a broad definition is satisfied, among others, by bacterial chemotaxis and phototropy of green plants (really treated as prototypes of intellectual activity [7, 11]) and by the actions of robots [18], which are usually not considered to be intelligent beyond robotics. Despite divergences from the more accurate descriptions of intelligence (numerous and often conflicting ones), this working definition agrees with most of its known implementations. This facilitates the main objective of the survey: to compare various objects and systems classified as intelligent in the literature and to identify their most common properties.

The dynamics of autonomous non-intelligent devices that flexibly perform their task in changing external conditions are described by the *theory of automatic control* (TAC) [14, 15, 22–24]. Within this theory, the evolution of a controlled object over time is guided by physical dynamic factors and control actions produced by an automatic control system (ACS) based on the observed parameters of the object (perhaps, noisy). The parameters characterizing the object's states at a given time instant, together with their inaccuracy estimates, are calculated from the information available to the system by recursive algorithms; their evolution in the *belief space* creates an approximate "image" of the factual dynamics in the *state space,* not precisely known to the control system.[4]

Thus, the ACS of a lifeless technical device already includes *the reflection of reality*, a phenomenon intrinsic to all generally recognized kinds of intelligence. Such reflection with statistically "blurred" estimates of the system state in the belief space has been further developed in *fuzzy control systems* [26], where the images of the automaton's environment directly include *information compression* (another inherent attribute of "real" intelligence). Both fundamental principles are practically implemented in environment mapping algorithms in robotics (*Simultaneous Location & Mapping* (SLAM) and other approaches, see the book [18]). In other words, some important signs of intellectual activity at the empirical level appear even in "proto-intelligent" automatic systems.

Indeed, TAC is a fascinating example of applying dynamic systems theory to the formulation and solution of control problems for various technical systems. Probably, one reason for its triumphant successes in the middle and late 20th century is the fact that many mechanical systems are adequately described by low-order differential or difference equations. Since the 1960s, many researchers have tried to translate the results of TAC to economic, social, and (or) living systems. The successes here, unfortunately, are rather modest. After all, see formulas (1a) and (1b), it would seem that we can apply the rich mathematical apparatus of TAC by saying "Let the controlled system be a linear discrete dynamic system." The problem is to justify the correctness of the notorious "let": to prove that such a description of the modeled economic, social, and (or) living system is adequate, that the system is identifiable, etc. Many researchers "forget" about such justifications, some of them fail on this way, but there are also exceptions (unfortunately, few and not massive).

[4] Within TAC, the agent's objective parameters are represented by the n-dimensional state vector $\mathbf{x} = (x_1, x_2, ..., x_n)$, whereas the information about these parameters available to the control system is represented by the m-dimensional observation vector $\mathbf{y} = (y_1, y_2, ..., y_m)$. The agent's dynamics in discrete time are reflected by the system of linear equations

$$\mathbf{x_{k+1}} = \mathrm{A}\mathbf{x}_k + \mathrm{B}\mathbf{u}_k + \omega_k , \quad (1a)$$

$$\mathbf{y_k} = \mathrm{C}\mathbf{x_k} + \theta_k , \quad (1b)$$

where $\mathbf{x}(t) \in \mathrm{X}$, $\mathrm{X} \subset R^n$ denotes the *state space*; $\mathbf{u}(t) \in U$ is the control vector; $U \subset R^p$ denotes the *control space*; $\mathbf{y}(t) \in Y$, $Y \subset R^m$ denotes the observation space (p, $m \le n$); $\mathrm{A} \in R^{n\times n}$ is the dynamic matrix; $B \in R^{n\times p}$ is the control matrix; $C \in R^{n\times m}$ is the observation matrix; $\omega(t)$ (system disturbance) and $\theta(t)$ (noise) are random variables; the subscripts k and $(k+1)$ correspond to sequential time instants (t, $t + \Delta t$). The state estimates $\{\hat{\mathbf{y}}_k\}$ and their variances $\{\mathrm{P}_k\}$ form the *belief space* $G=\{b_k\}=\{\hat{\mathbf{y}}_k, \mathrm{P}_k\}$, where, e.g.,

$$b_0 = (\hat{y}_0, P_0),\ b_{k+1} = K(b_k, \mathbf{u_k}, \mathbf{y_{k+1}}),\ \mathbf{u_{k+1}} = f(b_k) ,$$

and K is the Kalman filter; in the simplest one-dimensional case, $\hat{y}_{k+1} = Ky_{k+1} + (1 - K)(\hat{y}_k + u_k)$ [25].

According to the totality of heterogeneous definitions, the basic features of "real" (general) intelligence are the following simultaneous properties:

- information perception (reflection), processing, use, and transmission;
- autonomy, including purposeful behavior;
- knowledge retrieval, accumulation and analysis; learnability;
- abstraction, generalization, and generation of new knowledge;
- conscious autonomous goal-setting, self-awareness, creativity, emotions, and reflexion.

As has been demonstrated, the first two features manifest themselves at the level of automata whose goals are programmed (set exogenously). When increasing the levels of the cognitive capabilities of agents, their intellectual properties are modified and extended. In addition to the basic features listed above, there arise *autonomous goal-setting* and active learnability (starting from animals), the mental reflection of the external world and action schemes in it, called a cognitive map [27] in psychology (also starting from animals, but more relevant to humans), conscious construction of images of reality and reflexion (mainly characteristic of humans, but also of some aspects of AI [28]), and, further, the higher forms of intellectual creativity. In this survey, we will illustrate different "filling" of the features at different levels of intelligence, with all reservations on some uncertainty of this fundamental concept. Most manifestations of individual intelligence at the empirical level have analogies with AI implementations as well as with the DI of systems of interconnected agents, which can have very different intellectual abilities.

In the analysis below, we use the "naïve" (intuitive) idea of *information* as an ordered reflection of external influences on a single agent or on a multi-agent system. The reflection of reality as a characteristic of intelligence cannot be precisely separated from learnability and the ability to create new knowledge.[5] The first two features of "abstract" intelligence underlie most AI implementations. However, the empirically obvious ability to invent new things is not directly reproduced by mathematical models, despite the existing formal description of creative human activity; for example, see the paper [29] and references therein.

Insight is a characteristic feature of human intelligence and elements of animal cognitive activity associated with the solution of non-standard tasks. It represents a spontaneous arrival at the correct answer in a *problem situation*. Spontaneous problem solving, which allows bypassing some sequential logical inference stages with the obligatory concentration of an individual on solution search, is a firmly established result of psychological experiments [30, 31] with numerous heuristics for applied inventive purposes [32].

Despite the successful application of fuzzy heuristics, the problem of *information compression (concentration)* in the process of intellectual activity still has no general formal description. Numerous variants of similar external influences in *pattern recognition* form sets that cannot be analyzed by direct enumeration. For example, consider three parameters of a monochromatic sound (a simplified frequency spectrum of musical instruments) that are distinguishable by persons with musical education: pitch (five octaves, i.e., 60 variants of a note), loudness (six steps from *pp* to *ff*), and duration (six variants from integer to 1/32). Together with three sound extraction forms (staccato, nonlegato, and legato), they constitute $60 \times 6 \times 6 \times 3 = 6480$ combinations for one note. In this case, a short fragment of a musical text of 100 notes has much over 10^{100} variants (a *googol*). As in text and image recognition, information compression here is achieved by training an individual; as a result, the set of combinatorially possible variants is drastically reduced and transformed into a relatively small set of blocks, e.g., melodies, harmonies, or natural language words.

Individual human intelligence is the subject of most research in cognitive sciences; the above-mentioned empirical characteristics were established for it. They have been serving as a basis for developing artificial intelligence since the mid-20th century. In the works of biologists, numerous parallels were found between the cognitive abilities of animals and human intelligence; for some species, they include the ability to abstract and count, the use of intermediary languages, the formation of reality images in consciousness (with all reservations about the conditionality of this term when applied not to the consciousness of humans), and insight as the main decision mechanism in a problem situation [33]. Animal communities also have a social structure (sometimes very complex [34]) and are capable of cooperative actions. Nevertheless, we will further discuss the DI of structured systems on the examples of OSs in human society: they are better studied, more relevant, and are characterized by much more diverse intellectual behavior.

Note that individual human intelligence cannot form and exist outside the social environment, where it is an integral part of collective intelligence. Also, modern intellectual activity, both individual and collective, increasingly involves computers, robots, vision systems, and other elements of AI. Thus, all existing forms of intelligence are inextricably linked and

[5] Hereinafter, the notion of *knowledge* is also used in the everyday intuitive sense.

constitute the collective AI of human society. Technical "proto-intelligent" systems and computer implementations of AI (also not existing outside the society) significantly contribute to the structure of the DI of modern social systems and, therefore, should be briefly characterized.

2. MODELS OF ARTIFICIAL INTELLIGENCE

In first approximation, artificial intelligence as a scientific direction is divided into computational and logical ones. Figure 2 presents its most common implementations, not exhausting all kinds. According to the Artificial Intelligence Index Report 2022 (Stanford University, USA), the total annual number of publications on AI topics in the world increased from 162 thousand in 2010 to 334 thousand in 2021; see [3535][6]. Without attempting to analyze such a huge amount of information, we will list some areas of intelligence modeling that are closest to the objectives of this survey.

2.1. Artificial Neural Networks

The most popular (and fashionable) technical tool for reproducing some aspects of intellectual activity is the *artificial neural network* (ANN) [36]. The functioning of an ANN includes the *reflection* of the perceived "input" (images, texts, audio signals, etc.) and the *compression* of information into digital arrays during processing, i.e., the fundamental features of intelligence that are implemented at the level of ACSs and robots (see the discussion above). Here, the new feature of AI is *learnability*, i.e., obtaining a correct output on known input types.[7] The purpose of information processing by artificial neural networks is still set from the outside.

A main area of AI research is the *deep learning of* ANNs and other computer architectures with sequential information compression. The achievements in this area include face and object recognition from video surveillance data (computer vision), digitization of manuscripts and oral speech, compilation of logically coherent texts, machine translation, automatic creation of program codes, and many others. However, computer tools for modeling ANNs still lack an adequate general scientific description (a general theory of AI), and compressed information arrays at intermediate layers are vulnerable to small disturbances, including targeted attacks [37]. In this sense, neural networks repeat the fate of many fundamental technical achievements of mankind (steam engine, electronic circuits, aerial vehicles, etc.): their strict quantitative theory was developed only after their prolonged application. The issues of structure and functions of modern ANNs were discussed in the survey [38].

2.2. AI Based on Mathematical Logic

A special direction of cognitive sciences is devoted to modeling intelligence by mathematical logic methods [39]. Within this approach, external information is perceived by a computer program in the form of *propositions*, or logical statements, which can be true or false. Propositions establish connections between entities (objects that somehow reflect reality) according to natural language rules. Logical operations underlie *propositional calculus* in an infinite countable set of propositions (i.e., are used to determine their truth or falsity based on a given set of axioms). This opens the way to deriving new meaningful statements from true statements, in particular, to proving mathematical theorems "automatically." *Logic programming* is the most important component of this approach.

[6] Also, see Kaspar'yants, D., Review of Stanford University's Artificial Intelligence Index Report 2022, https://rdc.grfc.ru/2022/05/artificial_intelligence_index_report_2022/ (Accessed February 15, 2023.)

[7] In the ANN scheme as a simplified imitation of the network structure of brain neurons, an incoming signal (e.g., a pixel-scan image) is transmitted by elements of the input layer to the elements of the next (*hidden*) layer; depending on the ANN structure, these elements translate the signal to the output element or to the subsequent intermediate layer (in *recursive* networks, to the previous layers as well). The activation functions of the nodes, or *filters*, which determine their ability to transmit the input signal to the next node, can have a variety of stepwise forms. The analysis result is outputted in the form of an n-dimensional vector ($p_1, p_2, \ldots, p_n$) with the probabilities $\{p_i\}$ of attributing the input information array to one of the n specified types ($\sum p_i = 1$). When *training* an ANN, by analogy with the functioning of a network of brain neurons, the "conductances" of links between nodes are varied to achieve the correct identification of known types of the incoming signal. Further, the training sample of input information arrays is replaced by the analyzed array. In actively developed multilayer *convolutional* ANNs, signals entering the *convolutional layer* of artificial neurons and exiting from it to the next layer, as well as the sets of filter functions, are represented as ordered arrays of numbers, usually called tensors in IT. For example, an image of $N_1 \times N_2$ pixels with C color channels is given by an array $\mathbf{X}$ of dimensions $N_1^{(1)} \times N_2^{(1)} \times C$ with the normalized intensities $\{X_{ijk}\}$ of all color channels for each pixel; the set S of $m \times m$ filters with the same number of color channels is given by an array of dimensions $m \times m \times C \times S$ with components $\{F_{ijkp}\}$. The transformation results of the input data array,

$$Y(p, q, s) = \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{c} F(i, j, k, s) \times X(p+i-1, q+j-1, k),$$

form a tensor $\boldsymbol{Y}$ of dimensions $(N_1 - m + 1) \times (N_2 - m + 1) \times S$ with a smaller number of components, which is an incoming information array for the next layer. (For simplicity, the component subscripts are given in parentheses.)

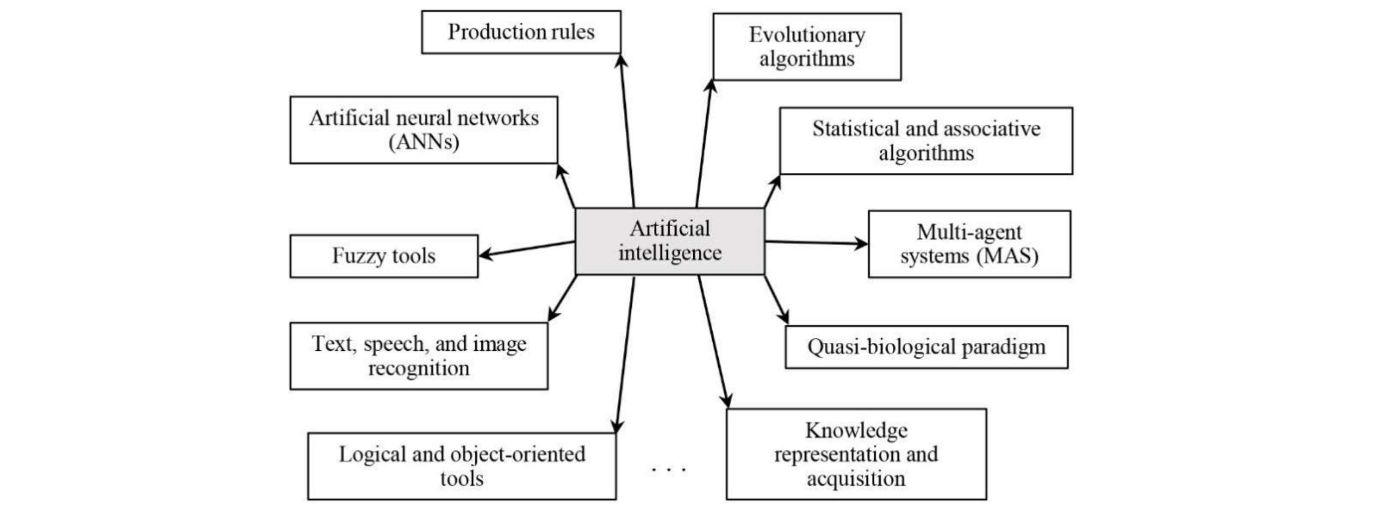

Fig. 2. The basic tools of mathematical and computer modeling of intelligence.

The implementation tools of AI, including different kinds of logic, were overviewed in detail in the book [40]. Mathematical logic methods have been embedded in a variety of computer tools facilitating intellectual tasks of users when processing large weakly ordered arrays of information in *knowledge bases* (as structured sets of established facts) of information, expert, and other "intelligent" systems. The approaches developed in this area include *Entity-Relationship* (ER) models and conceptual design [41], *semantic networks* [42], and other processing methods for weakly structured sets, as well as *fuzzy logic* (see Part B in the book [40]) and the modern promising direction of *Formal Concept Analysis* (FCA) [43].

The purely mathematical logic-based implementations of autonomous learning AI for creative tasks have not been described in the literature. Most of all, logical approaches are applied to generate and use knowledge based on computers and mathematical algorithms, particularly in combination with ANNs (see the survey [38]). Apart from pattern recognition, automatic theorem proving, knowledge generation [44], and superhuman-level program development for strategic games (chess, Go)[8], an important task here is *Natural Language Processing* (NLP), i.e., the representation, extraction, and handling of information expressed in a natural language. The greatest successes have been achieved in creating combined "human + computer" intelligent systems, where the dialog of the user with the program (input and output of information) is implemented in a natural language with simple graphics. However, this promising area is not the subject of our survey. Conceptual problems of artificial intelligence were discussed in the book [45].

Without disputing the merits of using mathematical logic in "intelligent" systems, we note that such systems cannot generate new nontrivial information except in the form of consistent inference from a given set of axioms and body of knowledge. Unlike ANNs, the "logical" implementations of AI are capable of creating new knowledge, manifesting another fundamental feature of intelligence (see above), but mathematical inference does not reproduce insight, the most essential mechanism of human cognitive activity. In addition, logical AI does not contain direct information compression schemes (unlike ANNs) and has only externally specified goal-setting (similar to ANNs). (Ontology design and use in knowledge extraction [1, 43, 44], as well as the development of self-learning computer programs, potentially capable of eliminating the most significant differences between AI and human intelligence, are still an area of intensive research.)

Another factor restricting the application of "automatic" logic is incomplete and often inconsistent data in knowledge bases, which is easily overcome by human thinking (see the book [45]). The biggest achievements here are also associated with the use of logical and semantic relations in the "human + computer" combination with human decision-making at each subsequent step. The matter concerns intelligent decision support systems (DSSs) based on the selection of related terms in the design of production, economic, and logistic schemes (cognitive maps in their mathematical sense [46]). This foundation is used to develop hybrid architectures of AI-supported human intelligence systems (see the Human-Computer Inter-

[8] According to O.P. Kuznetsov's reasonable comment, the very fact of equal competition between a grandmaster and a computer, excelling him (or her) millions of times by computational capabilities, testifies to the fundamental advantages of the information processing system in the human brain. Many modern R&D works in the field of AI seek to formalize the activity of human intelligence (see the book [45]).

action part in the book [21]), also not discussed in our survey.

A crucial and still little explored feature of human intelligence is the *intermittent* and asynchronous type of problem solving, contrasting with the existing continuous methods of its computer modeling. This feature manifests itself as *intermittent control* in a wide range of human activities: driving, work of operators, strategies of stock exchange players, and many other kinds of behavior [47]. All of them are characterized by the continuous monitoring of the situation in combination with episodic dynamics-correcting actions; in some cases, this combination resembles, but is not equivalent to, standard continuous control. Intermittent control is seemingly an evolutionarily selected means of saving brain resources: it is also observed in animals and is probably closely related to the phenomenon of insight. Some elements of intermittent control are evident in the DI of biological and social systems.

3. THE SIMPLEST FORMS OF DISTRIBUTED INTELLIGENCE

According to the aforesaid, the fundamental features of individual human intelligence (creating the images of the surrounding reality, recording and compressing external information, goal-setting, learnability, and "inventing new things") partially manifest themselves even in "proto-intelligent" automatic control systems and in computer implementations of AI (in a more complete form). Below we will see that different kinds of distributed intelligence in multi-agent systems exhibit analogous properties. Their implementation mechanisms in different kinds of DI also reveal great similarity. This suggests that the "intelligent" character of the system is determined by some fixed and common set of conditions. Following most definitions, we will treat the *property of emergence* as a sign of distributed (for groups of people, *collective*) intelligence [21]. (This property is the ability of a system of interacting agents to cope with tasks that exceed the capabilities of individual agents.)

By a *multi-agent social system* (MASS) we will understand a dynamic set of autonomous agents that interact with an environment and with each other. Each individual agent perceives and processes information coming from the environment and uses it to achieve some *goal* (an optimal result of its temporal evolution). Interacting biological entities of the same species (ants, bees, primates, and humans) form a social system in the narrow sense. Lifeless programmed agents (robots, drones, etc., see [7, 8]) form an *artificial* MASS in the presence of interaction between them.

The goal-setting of agents in most systems follows from their biological or social nature; for a system of interacting technical devices (or, e.g., an army unit), however, the goal is set from outside. Individual agents form an MASS of *level 1*. Interacting MASSs (economic entities, organizational systems, political parties, etc.) can themselves act as agents in higher-level social systems. The systems formed by human interactions and their activities were generally classified in the book [48].

Vast literature is devoted to modeling the dynamics of MASSs based on game theory [28, 49] and methods transferred from statistical physics ([7, 9, 50, 51], see Fig. 3). The dynamics of several systems (transportation flows [52], some economic processes [53], stock exchanges [54], influence [50, 55], and opinion spreading [56] in social environments) can be forecasted on short horizons using mathematical models. In such models, the processing and use of information by agents are usually considered indirectly in the form of quasi-physical potentials of repulsion of moving agents from obstacles and their attraction to targets [51, 52], striving to maximize utility [53, 55], clustering of like-minded people and ignoring social network users with opposing opinions [56], and other initially specified factors guiding collective dynamics.

However, the description of society is generally complicated by vaguely defined parameters, ambiguous dependencies, multilevel reflexion [28] and, on the other hand, by the boundedly rational behavior of agents [57, 58], as well as by other circumstances preferably considered at the verbal level in the humanities. It hampers calculations, formally possible after introducing quasi-physical "forces." As a rule, agent models based on the balance of quantitatively or qualitatively assessed factors (payoffs, influence, and interests) and maximization of an energy-like objective function under random disturbances (noise) provide only a qualitative description of social systems (see Fig. 3).

In our point of view, *distributed intelligence* as the ability of interconnected agents to perceive, process, and use information cooperatively [59] is the fundamental characteristic of MASSs that has incomplete reflection in the existing models. The system dynamics are not reducible to the simple sum of the evolutions of its agents; in particular, it includes system goal-setting. The *objective goal* of an MASS on a given time interval is to achieve an optimal state for the system as a whole. The formal optimality criterion (utilitarian, egalitarian, or another [53], Pareto optimum or Nash equilibrium, etc.) often depends on the preferences of researchers. Nevertheless, at the empirical level, it is difficult to deny that various systems, including those without the complex intellectual

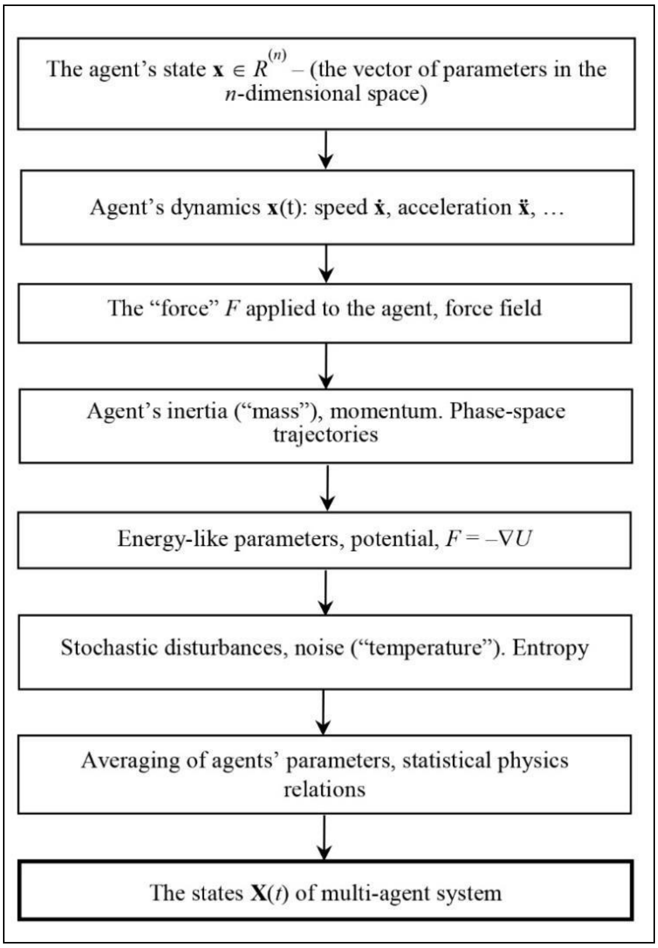

Fig. 3. The description of a multi-agent system by interdisciplinary physics methods.

activity of agents (a bee swarm, an ant colony, or a market) reproducibly and flexibly pursue systemic goals.[9] Essentially, the presence of individual consciousness in people is not a necessary condition for the existence of DI: systems with considerably different levels of the cognitive capabilities of agents can exhibit similar dynamics.

Distributed intelligence guiding system evolution should be directly considered in the description of social processes. Multiple kinds of DI, including distributed AI of technical systems, human-computer interaction schemes, biological MASSs, and various aspects of human activity were discussed in the monograph [21]. This survey is intended to analyze the conditions of DI manifestations in the dynamics of multi-agent systems (including those consisting of agents with small or zero cognitive capabilities) and then formulate general operation principles for an "intelligent agent."

[9] In most cases, objective goals include self-preservation of the system: preserving or increasing the number of agents and supporting their functioning.

3.1. Swarm Intelligence

Many autonomous agents forming a system can move independently (motor, self-moving, or living particles). Movement correlations arise when exceeding a certain threshold number of such particles in a unit volume (e.g., in flagellate bacteria). They gradually change the collective motion of particles from disordered (swarm) to vortex mode, also called bacterial turbulence [60]. Going back to the early 2000s, this term does not refer to turbulent fluid motion under a sufficiently high Reynolds number (Re~10^3): the size and velocity of flagellate bacteria (measured in μm and tens of μm/s, respectively) correspond to the values Re~10^{-3}–10^{-4}. On the usual macroscopic scale of distances and velocities, this matches the movement of massless particles in a very viscous medium [61]. However, due to hydrodynamic interaction, bacteria turn and are entrained behind neighboring moving particles, forming bacterial vortices (Fig. 4).

Movement correlation under a sufficiently high density of agents and a low noise level is implemented in all systems of self-moving particles [7]. It is also observed in lifeless systems of anisotropic colloidal particles and in hybrid materials, i.e., lyophilic liquid crystals with live bacteria added to them (see the paper [61]).

In biological MASSs of individuals capable of active information perception (insects, fish, and birds), there are "libraries" of different cooperative dynamics modes activated by external influences (e.g., flock takeoff in case of danger). The mode switching of the system in response to changes in external conditions serves as a means of adaptation that contributes to the survival of the species [7, 11]. This phenomenon, called *swarm intelligence* in cognitive sciences, has numerous analogies among other types of the DI of multi-agent systems [11].

Since the 1990s, various kinds of cooperative motion of living particles have been reproduced in simple multi-agent models (see the survey [7] and references therein). In a frequently used computer model (I. Cousin et al. [6]), the collective dynamics of a system of agents are determined by the ratio of the radii of three domains corresponding to the repulsion from nearest neighbors (R_1), the movement correlation (R_2), and the attraction to the group center (R_3). The modes of disorderly displacements within the group (*swarm*), vortex motion (*torus*), and coordinated motion (*flock*) are implemented depending on the density of agents, their velocities, and noise level. Hysteresis is observed when changing the radius of movement correlation (Fig. 5). In the real dynamics of a *school of fish*, movement correlations are affected by hydrodynamic interaction with neighboring individuals [62]. In large

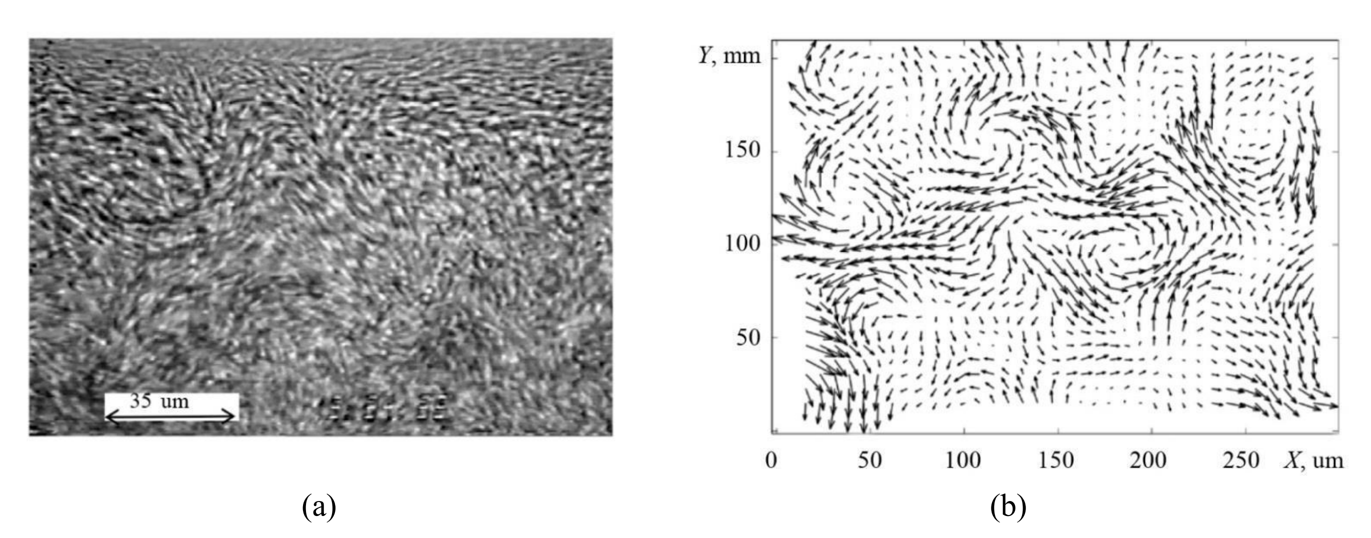

Fig. 4. The vortices of *Bacillus subtilis*: (a) photo and (b) velocity distribution [60].

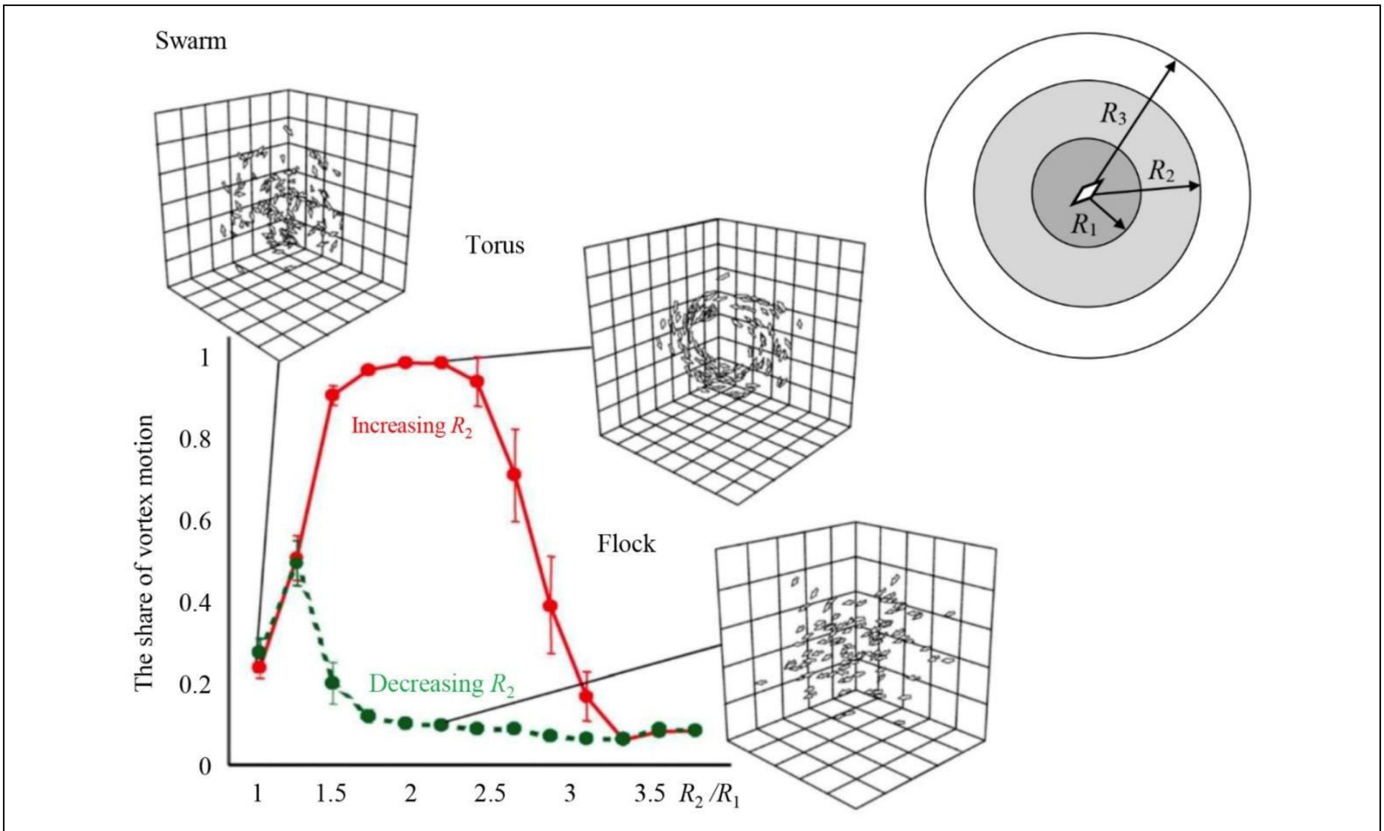

Fig. 5. The collective motion modes of agents depending on the ratio of the repulsion R_1 and movement correlation R_2 radii [6].

flocks of birds, where vortex structures are not formed, rich collective dynamics are created by the mechanism of *distributed switching leaders*, also reproduced in agent-based models: if there is a sufficient number of individuals equally changing the flight direction, the entire flock will correct its motion accordingly (Fig. 6).

The collective motion of biological MASSs always includes elements of chaotic behavior, especially pronounced in insects. The disordered arrangement and random movements of agents are an adaptive factor facilitating the recognition of danger and food sources. In particular, due to chaotic all-around view and movement correlations, a swarm of living particles performs collective surveillance far beyond the capabilities of its individuals [4–6, 11]. The dynamics of a school of fish attacked by a predator (avoidance response) were simulated, and complex structural changes were reproducibly captured under different attack strategies despite the complete parameterization of the agents (Fig. 7). Random changes in direction and velocity are also characteristic of birds in a flock; when several individuals deviate from a common course in the same way, they can become temporary

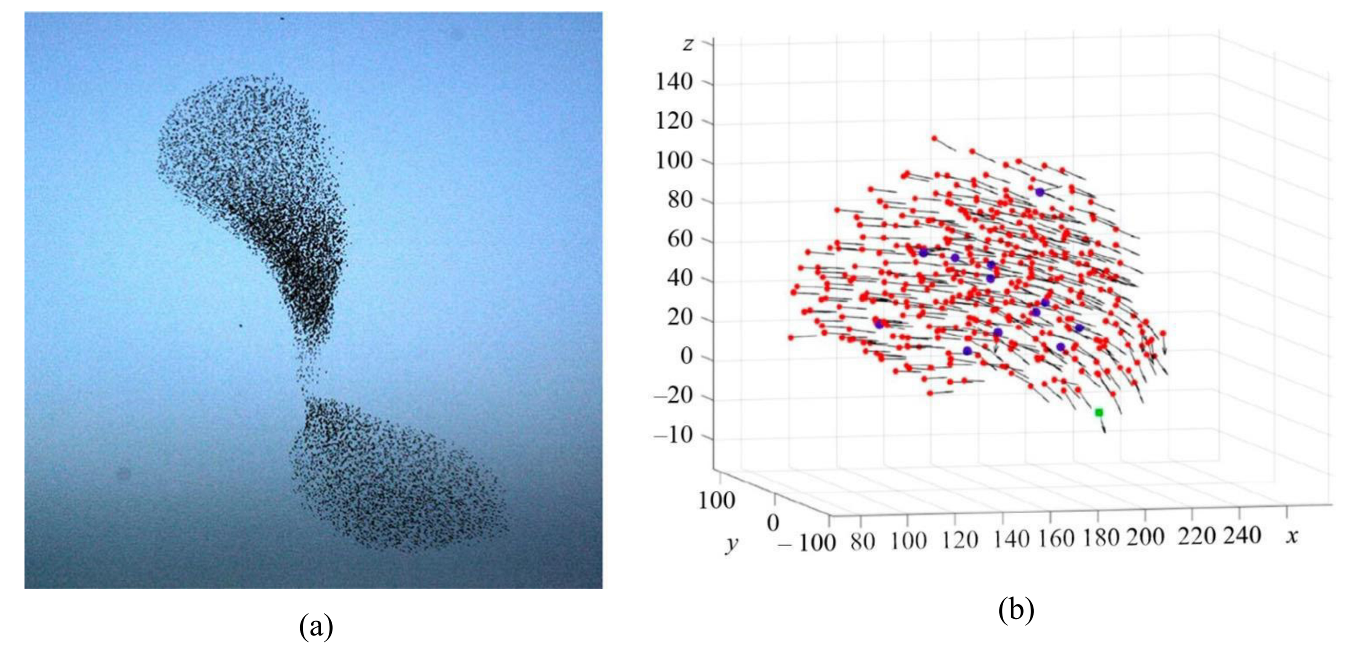

Fig. 6. Collective motion of birds: (a) a flock of starlings in the air and (b) movements of agents in the model. Blue color indicates temporary leaders [63].

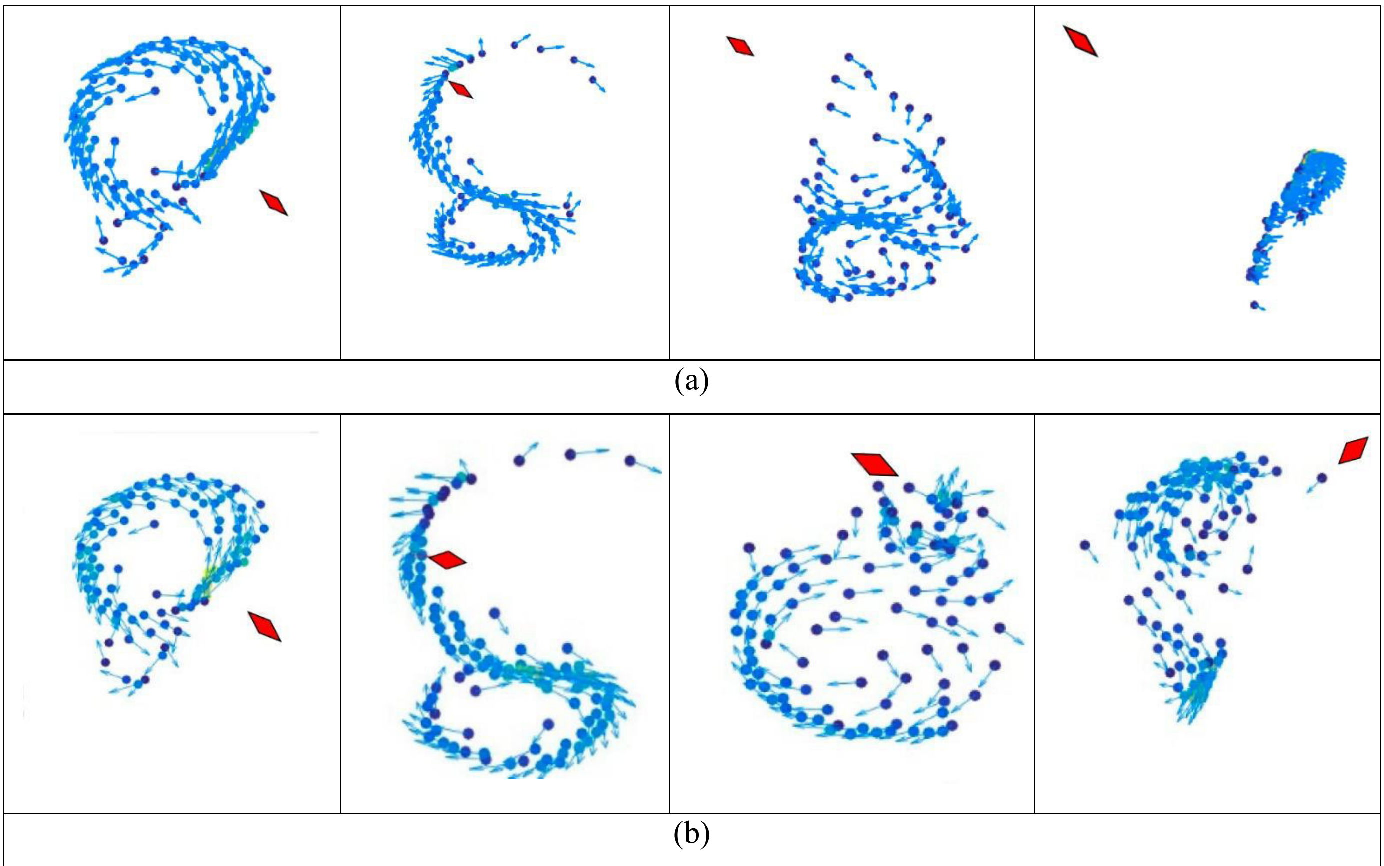

Fig. 7. Reorganizing the torus of moving agents under different predator strategies (shaded rhombus): (a) attack along a straight line and (b) attack on the nearest individual [62].

leaders influencing the overall flight path [63]. This ensures a flexible response of the system in a changing environment, which is not reduced to combinations of programmed actions. The mechanisms manifesting swarm intelligence (movement correlation, fluctuations of positions and velocities, and initiation of collective dynamics by the local majority of similarly moving individuals) have no general theoretical description but are successfully reproduced in simple agent models.

A peculiarity of social insects is "recording" information on the surrounding landscape using *pheromones*: chemical attractants recognized by other individuals within the system.[10] Ant paths labeled with pheromones, by a positive feedback mechanism, highlight optimal paths to food sources and other objects of ant colony activity (Fig. 8a) [7, 11].[11] A complex system of chemical insect behavior regulators forms a hierarchical system of ants, termites, and beehives [4]. Chaotic actions of individuals during construction and care for offspring are synchronized due to interactions within the system, generating cycles (Fig. 8b). Cooperative actions at the maximum of cycles are another adaptive mechanism selected by evolution to concentrate the colony's efforts, contributing to survival in unfavorable environments [6].

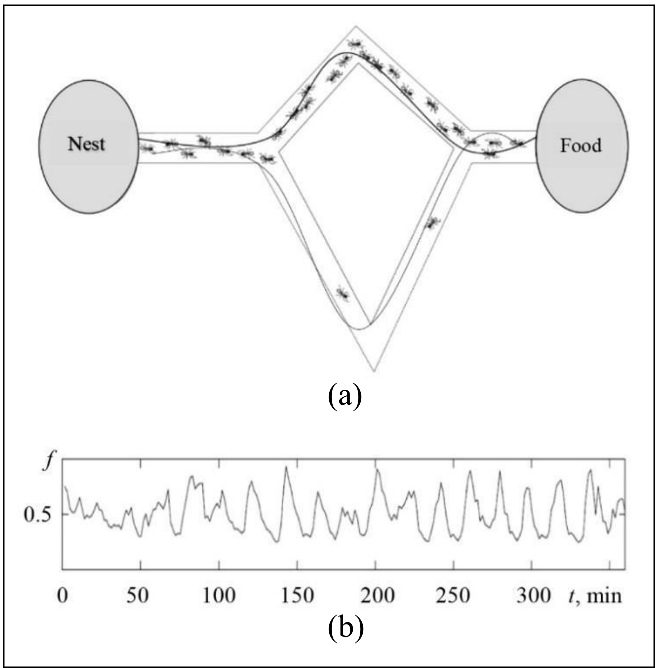

Fig. 8. Collective dynamics of social insects: (a) a short path to a food source chosen by ants [11] and (b) activity cycles (*f* denotes the share of active individuals) during nest construction [**6**].

[10] In a simpler example, the attack by wasps on a human approaching the nest is also initiated by their chemical secretions that increase aggressiveness [4].

[11] By a similar mechanism, with replacing olfaction by vision, footpaths appear on snow and grass (see [64]).

3.2. Swarm Intelligence in Robotics

Cooperative effects and dynamic structures also arise in artificial MASSs composed of autonomous interconnected vehicles that can perceive the positions of neighbor devices and avoid collisions when moving to a given target [8, 9]. Groups of robots can perform a common, externally assigned task with collective information processing using the distributed (not centralized) control mechanism. The swarm intelligence of systems of relatively simple and inexpensive interconnected technical devices (*swarm robotics*), in many respects imitating the DI of social insects, can execute several practically important operations:

– geometrical reconfiguration (making a *formation*) and collective motion;

– joint gathering and dispersal over a spatial area without loss of communication;

– functional separation;

– search, transportation, and joint movement of objects;

– collective positioning and mapping.

In robot formations, the "robustness" of collective actions is achieved: the system devices are interchangeable, and the system is insensitive to breakdowns of a certain number of elements [18, 65]. A typical applications-relevant example is the collective navigation of a group of flying drones [9]. In this direction of robotics, the terminology of "a social community of robots" has already emerged [11, 66], although its current tasks still refer to the optimization of joint motion.

3.3. Simulating Swarm Intelligence in Computer Algorithms

The idea of swarm intelligence is used in computer search and optimization algorithms, where increased computational power has changed both the technique and the computational strategy. In recent decades, several *Nature-Inspired Metaheuristics* (NIMs) *algorithms* have been developed in the field of information technology; in these algorithms, the simulation of swarm intelligence allows optimizing the solution of difficult computational problems[12] [40, 67,

[12] For example, consider the *ant colony* algorithm [67], first proposed in 1992 for finding the optimal path in complex graphs (*the traveling salesman problem*). In this algorithm, the transition probability of the representation point between vertices *i* and *j* is given by

$$P_{ij}(t)=\frac{\tau_{ij}(t)^{\alpha}\,d_{ij}(t)^{-\beta}}{\Sigma\tau_{ij}(t)^{\alpha}\,d_{ij}(t)^{-\beta}}$$

and depends on the path length d_{ij} between the vertices and the "pheromone level" τ_{ij} on this path. (Here, α and β are empirical

68]. The range of such computational algorithms is much wider than those of heuristics simulating swarms of living organisms (including the synchronized flashes of fireflies). In particular, the NIM domain covers schemes reproducing the social behavior of highly organized animals (wolves, lions, and monkeys [70]), as well as *genetic* and *cultural* algorithms [68]. Such computations involve changes in a population of virtual agents inheriting adaptive features due to random mutations (the genetic algorithm) or learning (the cultural algorithm). The rich variety of *natural computing* methods was considered in the survey [71].

Emergent effects assisting to achieve the optimum of an objective function in genetic algorithms provide a new look at the content of biological evolution. For example, in [59], the objective function of a simple evolutionary model (the time to pass a corridor with obstacles by a group of agents) was minimized in successive cycles by replacing the "slow" 50% of agents with copies of "fast" agents with fluctuations in their parameters (the acceleration when moving from left to right, repulsion from obstacles, and following neighbor agents within the radius of view). As a result of training, the corridor travel time of the last agent in the group decreased by 15–16% (Fig. 9a), and their chaotic drift with numerous collisions was replaced by a coordinated motion in clumps, which was not specified in the original model (Fig. 9b). Such results point to the analogy of evolutionary processes, in which the adaptability of biological species is optimized, with the development of their "systemic DI," long discussed at the qualitative level [72].

One direction of modern computer research is the artificial evolution of robots and their control programs with arbitrarily specified adaptive features [73]. Note that NIM algorithms are heuristic: they neither surely find the global optimum nor (as a rule) give estimates of how far the numerical solution will be from the exact one.

parameters, and the sum in the denominator corresponds to all possible routes from i to j). In the general case, the factor $d_{ij}^{-\beta}$ is replaced by the *heuristic value* $C_{ij}^{(k)}$, which is computed algorithmically for each partial solution (route). The pheromone level decreases at each discrete time step t and increases as new agents traverse the route:

$$\tau_{ij}(t+1) = (1-\varepsilon)\tau_{ij}(t)\sum_k \frac{\gamma}{C_{ij}^{(k)}}.$$

Here, ε and γ denote the evaporation coefficient and the value of the best solution, respectively; summation is performed over all routes $i \to j$ found.

By imitating the positive feedback of the attractiveness of real ant paths and the number of insects that have traveled along them, the ant colony algorithm economically finds suboptimal solutions of problems where computational cost grows factorially in the case of direct (straightforward) enumeration. In the alternative *bee colony* algorithm, each virtual agent randomly searches the best solution (e.g., the maximum of an objective function $U(p_1, p_2,..., p_n)$) along a connected trajectory in the space of parameters $\{p_i\}$ considering the common knowledge of the best point $\mathbf{p}^*(t)$ at each time step (the *blackboard*) [67]. In the general scheme of swarm-based optimization algorithms, the position and velocity of a virtual agent particle in the n-dimensional parameter space at the time step $(t + \Delta t)$,

$$\begin{aligned} x_i(t+1) &= x_i(t) + v_i(t)\Delta t, \\ v_i(t+\Delta t) &= \omega v_i(t) + a_l r_l\left[l_i(t) - x_i(t)\right]\Delta t \\ &+ a_g r_g\left[g(t) - x_i(t)\right]\Delta t, \end{aligned} \qquad (2)$$

are calculated from the position and velocity at the previous time step considering the position of the local optimum $l_i(t)$ in the agent's radius of view and the global optimum $g(t)$ from the data of all agents with empirical parameters ω (inertia), R (the radius of view), a_l and a_g (the acceleration coefficients to the optimums), and random numbers r_l, $r_g \in [0, 1]$. In this modification of the classical molecular dynamics method, the factors $a_l r_l[l_i(t) - x_i(t)]$ and $a_g r_g[g(t) - x_i(t)]$ at the time increment Δt in (2) are called the *cognitive factor* and *social factor*, respectively [69].

CONCLUSIONS

According to the analysis results, several basic features of intelligence as a fundamental phenomenon (autonomous perception and processing of external information, creation of images of the surrounding reality, goal-setting, learnability, and adaptation to changing external conditions) partially manifest themselves even in lifeless "proto-intelligent" automatic control systems as well as in different versions of computer artificial intelligence and in the collective dynamics of systems consisting of agents with small (social insects, fish, birds, etc.) or zero cognitive abilities (formations of robots and nature-inspired computer metaheuristics). A common feature of such systems is the emergent ability to perceive, process, and use external information far beyond the capabilities of individual agents.

The distributed intelligence of multi-agent systems in all its known forms reveals common features determined by the collective processing of information, not necessarily subjected to reflexion by the consciousness of agents. The simplest swarm modifications of DI in biological communities against the background of random behavior of individuals demonstrate selection and "memorization" of useful information, systemic goal-setting, and standard reactions to external influences (collective motion modes of fish and birds, ant paths, and cycles of synchronized dynamics of social insects). Elements of chaos are themselves an integral part of DI, creating the all-around view of a flock of fish, variable movements of a flock of birds, and other actions facilitating collective survival.

The capabilities of DI are determined by the intensity and structure of interactions between agents, their individual cognitive abilities, and the balance between

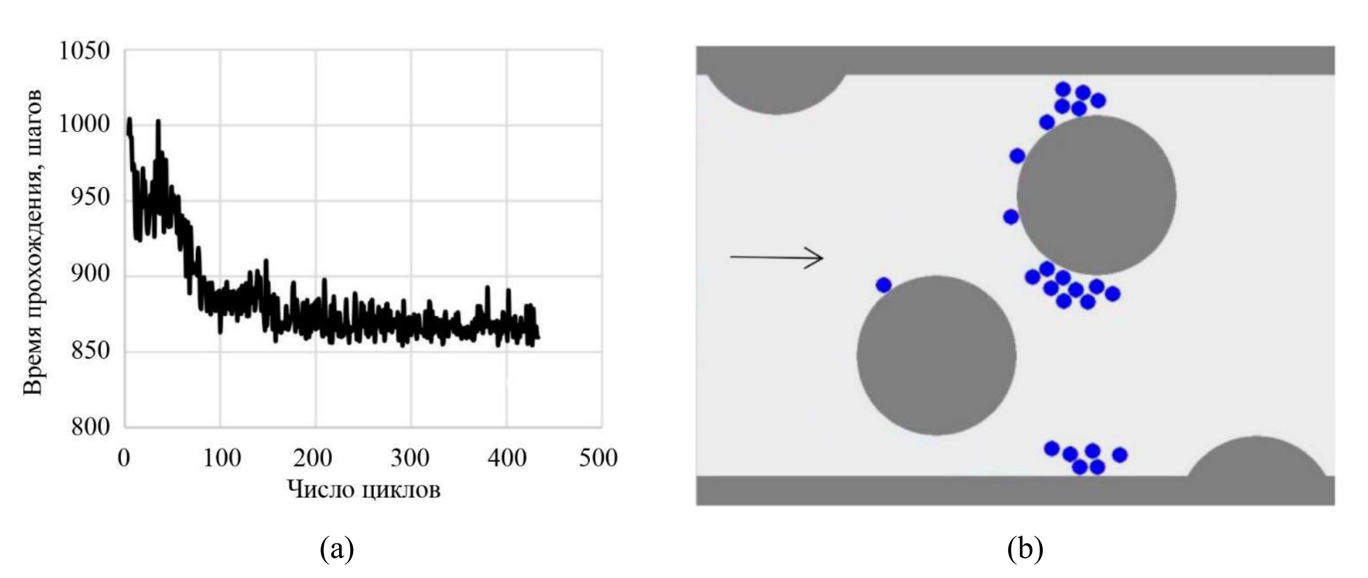

Fig. 9. Passing a corridor with an obstacle by agents in the genetic selection algorithm: (a) travel time (the number of steps) in successive cycles and (b) movements of agents in the "trained" system. The arrow indicates the direction of motion [59].

the degree of order of the system and random noise, which plays an important role in the dynamics of the system. Simulation of swarm intelligence in robotics and nature-inspired computer algorithms allows reproducing the emergent properties of DI (in particular, when searching suboptimal solutions of complex computational problems). The successful use of genetic and cultural computational algorithms in several applications opens up the possibility for a "cognitive" interpretation of the evolution of biological species, also revealing the essential features of DI. The discussion of more complex and effective forms of collective intelligence in social systems consisting of humans will be presented in part II of the survey.

Acknowledgments. *We are grateful to Prof., Dr. Sci. (Eng.) O.P. Kuznetsov, Cand. Sci. (Eng.) A.V. Makarenko, and Prof., Dr. Sci. (Eng.) P.O. Skobelev for careful reading of the manuscript and helpful remarks.*

REFERENCES

1. Novikov, D.A., *Cybernetics: From Past to Future*, Cham: Springer, 2016.
2. Cavens, D., Gloor, C., Illenberber, J., et al., Distributed Intelligence in Pedestrian Simulations, in *Pedestrian and Evacuation Dynamics 2005*, Waldau, N., Gattermann, P., Knoflacher, H., and Schreckenberg, M., Eds., Berlin–Heidelberg: Springer, 2007, pp. 201–212.
3. Fu, X., Gao, H., Cai, H., et al., How to Improve Urban Intelligent Traffic? A Case Study Using Traffic Signal Timing Optimization Model Based on Swarm Intelligence Algorithm, *Sensors*, 2021, vol. 21, art. no. 2631.
4. Kipyatkov, V.E., *Mir obschestvennykh nasekomykh* (The World of Social Insects), Leningrad: Leningrad State University, 1991. (In Russian.)
5. Chittka, L. and Rossi, N., Social Cognition in Insects, *Trends in Cognitive Sciences*, 2022, vol. 26, no. 7, pp. 578–592.
6. Couzin, I.D., Collective Cognition in Animal Groups, *Trends in Cognitive Sciences*, 2009, vol. 13, no. 1, pp. 36-43.
7. Vicsek, T. and Zafeiris, A., Collective Motion, *Phys. Rep.*, 2012, vol. 517, pp. 71–140.
8. Korepanov, V.O., *Modeli refleksivnogo gruppovogo povedeniya i upravleniya* (Models of Reflexive Group Behavior and Control), Moscow: Trapeznikov Institute of Control Sciences RAS, 2011. (In Russian.)
9. Vásárhelyi, G., Viragh, C., Somorjai, G., et al., Optimized Flocking of Autonomous Drones in Confined Environments, *Science Robotics*, 2018, vol. 3, no. 20, art. no. eaat3536.
10. Surowiecky, J., *The Wisdom of Crowds*, New York: Anchor Books, 2005.
11. Kennedy, J., Eberhart, R.C., and Shi, Y., *Swarm Intelligence*, San Francisco: Morgan Kaufman Publishers, 2001.
12. *The Cambridge Handbook of Intelligence*, 2nd ed., Sternberg, R.J., Ed., Cambridge: Cambridge University Press, 2020.
13. Kornhaber, M.L., The Theory of Multiple Intelligences, in *The Cambridge Handbook of Intelligence*, 2nd ed., Sternberg, R.J., Ed., Cambridge: Cambridge University Press, 2020,
14. Polyak, B.T., Khlebnikov, M.V., and Rapoport, L.B., *Matematicheskaya teoriya avtomaticheskogo upravleniya* (Mathematical Theory of Automatic Control), Moscow: LENAND, 2019. (In Russian.)
15. *Teoriya upravleniya (dopolnitel'nye glavy)* (Control Theory: Additional Chapters), Novikov, D.A., Ed., Moscow: LENAND, 2019. (In Russian.)
16. Elhoseny, M. and Yuan, X., Eds., *Artificial Intelligence Applications for Smart Societies. Recent Advances*, Studies in Distributed Intelligence, Cham: Springer, 2021.
17. Rzevski, G. and Skobelev, P., Emergent Intelligence in Large Scale Multi-agent Systems, *Intern. J. of Education and Inform. Technol.*, 2004, vol. 1, iss. 2, pp. 64–71.
18. *Springer Handbook of Robotics*, 2nd ed., Siciliano, B. and Khatib, O., Eds., Berlin: Springer-Verlag, 2015.
19. Legg, S. and Hutter, M., A Collection of Definitions of Intelligence, in *Adv. Artific. General Intelligence: Concepts, Archi-*

tectures and Algorithms, Goertzel, B. and Wang, P., Eds., Amsterdam: IOS Press, 2007, pp. 17–25.
20. Encyclopædia Britannica. URL: https://www.britannica.com/ (Accessed February 15, 2023.)
21. Malone, T.W. and Bernstein, M.S., *Handbook of Collective Intelligence*, Cambridge: MIT Press, 2015.
22. *Spravochnik po teorii avtomaticheskogo upravleniya* (Handbook of Automatic Control Theory), Krasovskii, A.A., Ed., Moscow: Nauka, 1987. (In Russian.)
23. Polyakov, K.Yu., *Osnovy teorii avtomaticheskogo upravleniya* (Foundations of Automatic Control Theory), St. Petersburg: St. Petersburg State Marine Technical University, 2012. (In Russian.)
24. *The Control Handbook*, vols. I–III, Levine, W., Ed., New York: CRC Press, 2010.
25. Vasile, C.-I., Leahy, K., Cristofalo, E., et al., Control in Belief Space with Temporal Logic Specifications, *Proc. IEEE Conference on Decision and Control*, Las Vegas, 2016, pp. 7419–7424.
26. Kudinov, Yu.I. and Pashchenko, F.F., Fuzzy Control Systems, in *Teoriya upravleniya (dopolnitel'nye glavy)* (Control Theory: Additional Chapters), Novikov, D.A., Ed., Moscow: LENAND, 2019, pp. 400–425. (In Russian.)
27. Tolmen, E.C., Cognitive Maps in Rats and Men, *The Psychological Review*, 1948, vol. 55, no. 4, pp. 189–208.
28. Novikov, D.A. and Chkhartishvili, A.G., *Reflexion and Control: Mathematical Models*, Leiden: CRC Press, 2014.
29. Belov, M.V. and Novikov, D.A., The Structure of Creative Activity, *Control Sciences*, 2021, no. 5, pp. 17–28.
30. Petukhov, V.V., *Psikhologiya myshleniya* (The Psychology of Thinking), Moscow: Moscow State University, 1987. (In Russian.)
31. Ponomarev, Ya.A., Aleksandrov, Yu.I., and Ushakov D.V., *Psikhologiya tvorchestva. Shkola Ya.A. Ponomareva* (The Psychology of Creative Activity. Ya.A. Ponomarev's School), Moscow: Institute of Psychology RAS, 2006. (In Russian.)
32. Altshuller, G., *And Suddenly the Inventor Appeared: TRIZ, the Theory of Inventive Problem Solving*, 2nd ed., Technical Innovation Center, 1996.
33. Zorina, Z.A. and Poletaeva, I.I., *Zoopsikhologiya. Elementarnoye myshleniye u zhivotnykh* (Zoopsychology. Elementary Thinking of Animals), Moscow: Aspekt Press, 2002. (In Russian.)
34. Tinbergen, N., *Social Behaviour in Animals with Special Reference to Vertebrates*, London–New York: Chapman & Hall, 1968.
35. *Artificial Intelligence Index Report*, Stanford University Human-Centered Artificial Intelligence, 2022, URL: https://aiindex.stanford.edu/report/ (Accessed February 15, 2023.)
36. Haykin, S., *Neural Networks: A Comprehensive Foundation*, 2nd ed., Singapore: Prentice Hall & Pearson Education, 1998.
37. Gu, T., Liu, K., Dolan-Gavitt, B., and Garg, S., BadNets: Evaluating Backdooring Attacks on Deep Neural Networks, *IEEE Access*, 2019, vol. 7, pp. 47230–47244.
38. Makarenko, A.V., Deep Neural Networks: Origins, Development, Current Status, *Control Sciences*, 2020, no. 2, pp. 3–19. (In Russian.)
39. Thayse, A., Gribomont, P., Louis, G., et al., *Approche logique de l'intelligence artificielle. T. 1: De la logique classique a la programmation logique*, Paris: Bordas Editions, 1988. (In French.)
40. *Springer Handbook of Computational Intelligence*, Kacprzyk, J. and Pedrych, W., Eds., Berlin: Springer, 2015.
41. Batini, C., Ceri, S., and Navathe, S.B., *Conceptual Database Design: An Entity-Relationship Approach*, Redwood: Benjamin/Cummings, 1992.
42. Lehmann, F., *Semantic Networks in Artificial Intelligence*, New York: Pergamon, 1992.
43. Poelmans, J., Ignatov, D.I., Kuznetsov, S.O., and Dedene, G., Formal Concept Analysis in Knowledge Processing: A Survey on Applications, *Expert Systems with Applications*, 2013, vol. 40, no. 16, pp. 6538–6560.
44. Vassiliev, S.N. Interactive Generation of New Knowledge Based on Automatic Means of Logical Inference, *Ontology of Designing*, 2023, vol. 13, no. 1, pp. 10–29. (In Russian.)
45. Kuznetsov, O.P., *Izbrannye trudy. Avtomaty, yazyki i iskusstvennyi intellect* (Selected Proceedings. Automata, Languages, and Artificial Intelligence), Moscow: Trapeznikov Institute of Control Sciences RAS, 2016. (In Russian.)
46. Kulinich, A.A., Semiotic Cognitive Maps. Part 1. Cognitive and Semiotic Approaches in Informatics and Control Sciences, *Control Sciences*, 2016, no. 1, pp. 2–10; P. 2. The Basic Definitions and Algorithms, *Control Sciences*, 2016, no. 2, pp. 24–40. (In Russian.)
47. Gawthrop, P., Loram, I., Lakie, M., and Gollee, H., Intermittent Control: A Computational Theory of Human Control, *Biological Cybernetics*, 2011, vol. 104, pp. 31–51.
48. Belov, M.V. and Novikov, D.A., *Modeli deyatel'nosti (osnovy matematicheskoi teorii deyatel'nosti)* (Models of Activity. Foundations of Mathematical Theory of Activity), Moscow: LENAND, 2021. (In Russian.)
49. Novikov, D.A., Models of Informational Confrontation in Mob Control, *Control Sciences*, 2015, no. 3, pp. 29–39. (In Russian.)
50. Chkhartishvili, A.G., Gubanov, D.A., and Novikov, D.A., *Social Networks: Models of Information Influence, Control and Confrontation*, Heidelberg: Springer, 2019.
51. Slovokhotov, Y.L. Physics vs. Sociophysics, *Control Sciences*, 2012, no. 1, pp. 2–20; no. 2, pp. 2–31; no. 3, pp. 2–34. (In Russian.)
52. Gasnikov, A.V., *Vvedenie v matematicheskoe modelirovanie transportnykh potokov* (Introduction to Mathematical Modeling of Traffic Flows), Moscow: Moscow Center for Continuous Mathematical Education, 2013. (In Russian.)
53. Romanovskii, M.Yu. and Romanovskii, Yu.M., *Matematicheskie nachala ekonofiziki* (Mathematical Foundations of Econophysics), Moscow–Izhevsk: Institute of Computer Studies, 2019. (In Russian.)
54. Shiryaev, V.I., *Finansovye rynki. Neironnye seti, khaos i nelineinaya dinamika* (Financial Markets. Neural Networks, Chaos and Nonlinear Dynamics), 5th ed., Moscow: LIBROKOM, 2013. (In Russian.)
55. Barabanov, I.N. and Novikov, D.A., Continuous-Time Dynamic Models of Mob Excitation, *Large-Scale Systems Control*, 2016, no. 67, pp. 71–86. (In Russian.)
56. Kozitsin, I.V., A General Framework to Link Theory and Empirics in Opinion Formation Models, *Sci. Rep.*, 2022, vol. 12, no. 5543, pp. 1–18.
57. Spiegler, R., *Bounded Rationality and Industrial Organization*, Oxford: Oxford University Press, 2014.

58. Novikov, D.A., Bounded Rationality and Control, *Autom Remote Control*, 2022, vol. 83, no. 6, pp. 990–1009.
59. Slovokhotov, Y.L. and Neretin, I.S., Towards Constructing a Modular Model of Distributed Intelligence, *Programming & Computer Software*, 2018, vol. 44, no. 6, pp. 499–507.
60. Dombrowski, C., Cisneros, L., Chatkaew, S., et al., Self-Concentration and Large-Scale Coherence in Bacterial Dynamics, *Phys. Rev. Lett.*, 2004, vol. 93, no. 9, pp. 098103-1–098103-54.
61. Aranson, I.S., Active Colloids, *Physics-Uspekhi*, 2013, vol. vol. 56, no. 1, pp. 79–92.
62. Deng, J. and Liu, D., Spontaneous Response of a Self-Organized Fish School to a Predator, *Bioinspiration & Biomimetics*, 2021, vol. 16, no. 4, art. no. 046013.
63. Cristiani, E., Menci, M., Papi, M., and Brafman, L., An All-Leader Agent-Based Model for Turning and Flocking Birds, *J. Math. Biology*, 2021, vol. 83, art. no. 45.
64. Helbing, D., Keltsch, J., and Molnar, P., Modelling the Evolution of Human Trail Systems, *Nature*, 1997, vol. 388, no. 3, pp. 47–50.
65. El Zoghby, N., Loscri, V., Natalizio, E., and Cherfaoui, V., Robot Cooperation and Swarm Intelligence, in *Wireless Sensor and Robot Networks: from Topology Control to Communication Aspects*, Singapore: World Scientific, 2014, pp. 168–201.
66. Kulinich, A.A., Karpov, V.E., and Karpova, I.I., *Sotsial'nye soobshchestva robotov* (Social Communities of Robots), Moscow: LENAND, 2019. (In Russian.)
67. McConnell, J.J., *Analysis of Algorithms: An Active Learning Approach*, Boston: Jones & Bartlett Publishers, 2001.
68. *Nature-Inspired Metaheuristic Algorithms for Engineering Optimization Applications*, Springer Tracts in Nature-Inspired Computing, Carbar, S., Toktas, A., and Ustin, D., Cham: Springer, 2021.
69. Winklerova, Z., Maturity of the Particle Swarm as a Metric for Measuring the Particle Swarm Intelligence, *Proc. 4th Int. Conf. Adv. Swarm Intel.*, Harbin, 2013, pp. 40–54.
70. Chakraborty, A. and Kar, A.K., Swarm Intelligence: A Review of Algorithms, in *Nature-Inspired Computing and Optimization: Theory and Applications*, Patnaik, S., Yang, X.-S., Nakamatsu, K., Cham: Springer, 2017, pp. 475–494.
71. Shiroky, A.A. and Kalashnikov, A.O., Natural Computing with Application to Risk Management in Complex Systems, *Control Sciences*, 2021, no. 4, pp. 2–17.
72. Dawkins, R., The Selfish Gene, Oxford: Oxford University Press, 1976.
73. Alattas, R.J., Patel, S., and Sobh, T.M., Evolutionary Modular Robotics: Survey and Analysis, *J. Intel. & Robotic Syst.*, 2019, vol. 95, pp. 815–828.

This paper was recommended for publication by V.E. Lepskiy, a member of the Editorial Board.

Received March 6, 2023,
and revised May 11, 2023.
Accepted May 22, 2023.

Author information

Slovokhotov, Yuri Leonidovich. Dr. Sci. (Chem.), Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia; Department of Materials Science, Moscow State University
✉ yurislovo@yandex.ru
ORCID iD: https://orcid.org/0000-0002-6669-6210

Novikov, Dmitry Aleksandrovich. Academician, Russian Academy of Sciences; Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
✉ novikov@ipu.ru
ORCID iD: https://orcid.org/0000-0002-9314-3304

Cite this paper

Slovokhotov, Yu.L., Novikov, D.A., Distributed Intelligence of Multi-Agent Systems. Part I: Basic Features and Simple Forms. *Control Sciences* **5,** 2–17 (2023). http://doi.org/10.25728/cs.2023.5.1

Translated into English by *Alexander Yu. Mazurov*,
Cand. Sci. (Phys.–Math.),
Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
✉ alexander.mazurov08@gmail.com

DOI: http://doi.org/10.25728/cs.2023.5.2

COURNOT OLIGOPOLY: STRATEGY CHOICE UNDER UNCERTAINTY AND OTHER PROBLEMS[1]

A.S. Shvedov

National Research University Higher School of Economics, Moscow, Russia

✉ ashvedov@hse.ru

Abstract. Firms operating in a market economy naturally strive to increase revenues. When large firms affect prices by their actions, this task involves nontrivial mathematics, i.e., game-theoretic oligopoly models. The survey is more concerned with Cournot competition than with Bertrand competition. The existence, uniqueness, and stability of Cournot equilibrium are discussed. The other issues under consideration are as follows: the entry of new firms into the market; the barriers that can be imposed for this; and the impact of such an entry on society's welfare as well as on total surplus and consumer surplus. The problems of collusion between firms are touched upon. Publications comparing the prices of goods, the profits of firms, and society's welfare under Cournot and Bertrand competition are overviewed. Much attention is paid to the problems faced by firms due to the ignorance of some current or future market conditions and the existing uncertainty. The issues of information sharing among firms are considered. One approach to reducing marginal cost is the purchase of licenses; licensing in a Cournot duopoly is also described. Computational methods for Cournot equilibria in the case of multi-product firms are presented. Finally, publications with particular applications of Cournot equilibria are considered.

Keywords: Cournot equilibrium, social efficiency, Bertrand equilibrium, information sharing, uncertainty, licensing, cartel formation, complementarity problem.

INTRODUCTION

If large firms operate in an industry, the prices of goods are determined not only by demand and production costs but also by the strategies of producers. The theory of oligopoly plays an important role when forming the strategies of firms. According to the classification in the book [1], the interactions occurring in an industry with a small number of firms can be quantity competition (firms determine their outputs), price competition (firms determine the prices of their products), or collusion. The interaction where all firms simultaneously choose their outputs, trying to predict the outputs of other firms, is called Cournot competition. The interaction where all firms simultaneously establish the prices for their products, trying to predict the prices of other firms, is called Bertrand competition. Cournot competition and Bertrand competition were also discussed in the book [2] (as Cournot oligopoly and Bertrand oligopoly, respectively). Cournot competition-based approaches may be preferred when outputs must be determined long before production. Information sharing among firms, e.g., costs and market demand, (or its absence) is essential. According to the paper [3], advertising and R&D expenditures are other elements of firms' strategies in addition to outputs and prices; this makes the model algebraically more complex but does not change it completely. Of course, mathematical modeling of non-simultaneous decision-making by firms is of certain interest, but such models will not be addressed below.

The classical Cournot competition model is as follows. Assume that n firms produce a homogeneous good sold at a uniform price. If L_i is the *capacity* of firm i, then its strategy is to produce a *quantity* (*output*) q_i, where $0 \le q_i \le L_i$, $i = 1, ..., n$. By assumption, the firm's *costs* C_i depend only on the output q_i

[1] This work was supported by Academician Nikolai Fedorenko International Scientific Foundation for Economic Research, project no. 2022-139.

whereas the *price of the* good P depends only on the *total output* $Q=\sum_{i=1}^{n} q_i$. Then the profit of firm i is given by

$$\pi_i(q_i, Q) = P(Q)q_i - C_i(q_i). \quad (1)$$

The function $P(Q)$ is monotonically decreasing, and each function $C_i(q_i)$ is monotonically increasing (except for the models with constant costs). The function $P(Q)$ is called the *inverse demand function* and the functions $C_i(q_i)$ are called *cost functions*. The objective of firm i is to maximize the profit π_i. In modern terms, this problem is a noncooperative game. (Note that Cournot's and Bertrand's works date back to the 19th century.) It is required to find a Nash equilibrium in the class of pure strategies profiles $q_1,...,q_n$. When applied to this problem, Nash equilibrium is often called *Cournot equilibrium* or Cournot–Nash equilibrium. A more general concept of equilibrium (with collusion between firms regarding their outputs) was discussed, e.g., in [3]. As it turns out, Cournot equilibrium is a special case of this equilibrium.

The problem formulated above motivated the appearance of numerous interesting and important mathematical works. However, Cournot's goal was to describe the economic realities lying between monopoly and perfect competition, in particular, to answer the following question: Does an increase in the number of firms match public interests? Among other achievements, scientists of the 19th century recognized that it is impossible to answer such questions without applying mathematical methods and solving optimization problems. But what is the level of mathematical sophistication required? According to some authors, search engines provide over 50 000 references for the query "Cournot equilibrium." In some papers, the presentation was limited to the formulation and proof of mathematical theorems; in others, economic conclusions were drawn or specific economic problems were considered. Undoubtedly, it would be wrong to claim that some link in the "theoretical mathematics–applied mathematics–particular applications" chain is more important than others. In addition, with such voluminous literature on the subject, not mentioning some publications in this survey does not mean their insignificance from the author's point of view.

The rapid development of Cournot oligopoly as a mathematical discipline began in the late 1950s–early 1960s. First, the existence and uniqueness of Cournot equilibrium were studied. Second, dynamic models were considered, and the stability of Cournot equilibria was analyzed. These problems are presented in Section 1 of the survey. Note that the aim of this paper is not to overview dynamic models of Cournot competition: the focus is on static models. However, at the initial evolution stage of Cournot oligopoly as a mathematical discipline, the issues of existence, uniqueness, and stability were considered in close connection.

Is the entry of a new firm into the market always desirable in terms of public interests? May such an entry reduce the output of existing firms and increase the price of goods, becoming therefore profitable for the entering firm, but decrease society's welfare? Should the authorities impose entry barriers for new firms, and what should these barriers be? How will these barriers affect total surplus and consumer surplus? The questions listed above are interrelated. Relevant publications are overviewed in Section 2. Many authors compared Cournot and Bertrand competition in terms of equilibrium prices, outputs, society's welfare, etc. These studies are presented in Section 3. To a large extent, the survey concerns probability theory methods when applied to the interaction of large firms. Section 4 is devoted to the works on strategy choice under uncertainty and information sharing among firms. Also, this section describes works on the sale of licenses. Cartel formation is discussed in Section 5. Finally, Section 6 considers the practical calculation of Cournot equilibria and their use in particular applications.

1. THE EXISTENCE, UNIQUENESS, AND STABILITY OF COURNOT EQUILIBRIA

Leaving aside the existence and uniqueness of Cournot equilibria for a while, we begin with the paper [4], which analyzed the stability of such equilibria. That paper considered a discrete-time model as follows. At each time step, by determining its output, each firm maximizes its profit under a linear inverse demand function $P(Q)$ known to all firms. Each firm assumes that the outputs of other firms will not change. The stability of the resulting system of difference equations was investigated. As it turned out, the solution is stable in the case of 2 firms only; if the number of firms exceeds 3, the solution becomes unstable.

Note that by that time, there was considerable literature on the stability of Walrasian equilibria; it was well known that the (in)stability of the solution depends on the adjustment process used in the model (e.g., see [5, *p. 643*]). The judgmental and undoubted practical applicability of the conclusions of [4], and the fact that the adjustment process in that work is chosen arbitrarily, stimulated further studies of the stability of Cournot equilibria. For a wider class of adjustment processes, it was shown in [6] that Cournot equilibrium can be stable for any number of firms.

The author [7] considered continuous-time models with different cost functions of firms and a nonlinear inverse demand function. Lyapunov functions were used to prove the stability of Cournot equilibria. The paper [8] combined the approaches from [7] and [5] to formulate general conditions for adjustment processes under which Cournot equilibria are stable. The researcher [9] relaxed the assumption that all firms produce a homogeneous good, but each firm produces only one good. (Such models are called differentiated goods models, in contrast to models where each firm produces multiple goods.) The stability of Cournot equilibrium was proved under "a sufficiently weak link between goods." The stability results from [7] were extended in [10] using the concept of conjectural variations. (For more details on conjectural variations, we refer to, e.g., [11].) Further results on the stability of Cournot equilibria were overviewed in [12–14].

In the case of firms producing a homogeneous good, the existence of Cournot equilibria was considered in [15] as follows. Assume that the functions $P(Q)$ and $C_i(q_i)$ are continuous (but not necessarily differentiable) and the function $\pi_i(q_i, Q)$ is concave in the argument q_i. Under these conditions, the existence of Cournot equilibrium was established. In addition, previous works containing special cases of the corresponding theorem were cited. The uniqueness theorem of Cournot equilibrium was proved in [16]. Also, we should mention the work [17] on concave noncooperative games: the results presented therein can be used to prove the existence and uniqueness of Cournot equilibria, but Cournot oligopoly was not considered as an example in that paper. The authors [15] and [17] showed the existence of equilibria using Kakutani's fixed point theorem. A simpler proof of the existence and uniqueness of Cournot equilibria for differentiable functions $P(Q)$ and $C_i(q_i)$ was provided in [18]. According to the counterexamples in the book [19, *pp. 4*, *5*], there may exist no Cournot equilibrium at all or there may exist several (nonunique) Cournot equilibria. The existence, uniqueness, and stability of Cournot equilibria can be considered with respect to firms producing multiple goods as well (e.g., see [20]). These issues were also covered in the book [19]. The existence of Cournot equilibria for the case of biconcave inverse demand functions was established in [21].

In a number of works, the economic processes under consideration were studied along the path of increasing the complexity of the mathematical tools involved. In the paper [22], the inverse demand function was understood as a multivalued mapping. In [23], the goal was to create a unified conceptual approach to Walrasian and Cournot equilibria as follows. Initially, the notion of economy was introduced in a way common for Walrasian equilibria with production. Cournot production was then defined as some probability measure, and Cournot equilibrium was understood as an equilibrium in the class of mixed (not pure) strategies. Other mathematical tools are also adopted to study Cournot oligopoly; for example, in [24], the theory of attractors was used to compare Walrasian and Cournot equilibria within discrete-time models.

The publication [25] considered firms producing a homogeneous good with different cost functions. The existence of Cournot equilibria was investigated. The focus was on the following question: how far can the previous conditions on the inverse demand function and cost functions (see [15, 18] and others) be relaxed without violating the existence theorem of Cournot equilibria? The results of that work partially overlap with those obtained independently in [26]. The paper [27] belongs to the same line of research. Subsequent publications on the subject were overviewed in [28] for the inverse demand function $P(Q) = a - bQ^{\beta}$, where β can be either positive or negative; the existence, uniqueness, and stability of Cournot equilibria were studied. The author [29] studied the effect of risk aversion on the strategies of firms.

Many researchers developed the classical Cournot competition model with application to certain economic problems. In [30], oil production was studied, and the profits of firms were maximized over a long time interval with discounting. Within the model described in [31], firms choose which labor (as a factor of production) to use, and their outputs are determined by their choice. According to [32], a mixed oligopoly is an oligopoly with one state-owned firm and several private firms. The state-owned firm seeks to maximize society's welfare, whereas private firms maximize profit (all firms produce a homogeneous good). The conditions were formulated under which Cournot equilibrium exists and is unique in such a mixed oligopoly. The effect of taxes on equilibrium outputs in a Cournot oligopoly was discussed in [33, 34]. In [35], the existence of Cournot equilibrium was proved under fixed cap prices. In [36], random yield models with Cournot competition were considered: the output of each firm has the form $q_i = \xi_i \bar{q}_i$, $i = 1, ..., n$, where $\bar{q}_i$ is the target quantity and ξ_i is a random variable. Total output, consumer surplus, and the entry of new firms into the market were investigated.

Correlated equilibrium (by definition, some probability measure) is a generalization of Nash equilibrium in pure strategies. The following result was established in [37, 38] under appropriate conditions: if there exists a unique Cournot equilibrium, it will also be a unique correlated equilibrium.

In several publications, reflexive game theory methods [39] were used to study Cournot competition. The conditions presented in [40] ensure convergence to a Cournot equilibrium when each firm gives the best response (in terms of maximizing its profit) to the outputs of other firms. In the case of one Leader and several Follower firms, the convergence to a Cournot equilibrium was discussed in [41]. Some examples of nonunique Cournot equilibria were given in [42]. Efficiency issues for static and dynamic games under different organizational modes of firms were studied in [43].

2. MARKET ENTRY AND EFFICIENCY ISSUES

Let the inverse demand function have the form $P(Q)=a-bQ$, where $a>0$ and $b>0$; let the cost functions have the form $C_i(q_i)=c_iq_i$, where $c_i\geq 0$. Assume that the capacity of each firm is unbounded. The values c_i, key for the subsequent analysis, are called *marginal cost* or unit production cost. Then, in accordance with formula (1),

$$\pi_i(q_1,\dots,q_n)=\left(a-b\sum_{j=1}^{n}q_j\right)q_i-c_iq_i$$

and

$$\frac{\partial\pi_i}{\partial q_i}(q_1,\dots,q_n)=a-2bq_i-b\sum_{j\neq i}q_j-c_i.$$

A strategy profile $q_1^*,\dots,q_n^*$ is a Nash equilibrium if

$$\pi_i(q_1^*,\dots,q_{i-1}^*,q_i^*,q_{i+1}^*,\dots,q_n^*)$$
$$=\max_{q_i}\pi_i(q_1^*,\dots,q_{i-1}^*,q_i,q_{i+1}^*,\dots,q_n^*)$$

for all $i=1,\dots,n$. The necessary condition of the maximum is the zero value of all partial derivatives:

$$\frac{\partial\pi_i}{\partial q_i}(q_1^*,\dots,q_{i-1}^*,q_i^*,q_{i+1}^*,\dots,q_n^*)=0$$

or

$$bq_i^*=a-c_i-b\sum_{j=1}^{n}q_j^*,\ i=1,\dots,n.$$

Summing the last n equations gives

$$Q^*=\frac{n}{b(n+1)}(a-\bar{c}),\qquad(2)$$

where $\bar{c}=\frac{1}{n}\sum_{i=1}^{n}c_i$ is the average marginal cost. By assumption, $a>\bar{c}$. Then

$$P(Q^*)=a-\frac{n}{n+1}(a-\bar{c}).\qquad(3)$$

The values q_i^* and $\pi_i(q_1^*,\dots,q_n^*)$ can be easily calculated:

$$q_i^*=\frac{1}{b(n+1)}\left(a-\bar{c}+(n+1)(\bar{c}-c_i)\right),\qquad(4)$$

$$\pi_i(q_1^*,\dots,q_n^*)=\frac{1}{b(n+1)^2}\left(a-\bar{c}+(n+1)(\bar{c}-c_i)\right)^2.\qquad(5)$$

The point q_i^* is the maximum point of the profit function $\pi_i(q_1^*,\dots,q_{i-1}^*,q_i,q_{i+1}^*,\dots,q_n^*)$ since π_i is concave in the argument q_i. In the case $c_i>\bar{c}$, the problem of negative (zero) outputs q_i^* may arise. This problem will be considered in detail in Section 6. Here, we suppose that $c_i<\bar{c}+(a-\bar{c})/(n+1)$ for all i. Note that the share of the output of firm i in total output is easy to calculate as well:

$$\frac{q_i^*}{Q^*}=\frac{1}{n}\,\frac{a-c_i}{a-\bar{c}}+\frac{\bar{c}-c_i}{a-\bar{c}}.$$

Due to formulas (4) and (5), if n is large enough, the difference ($\bar{c}-c_i$) can play a greater role than the difference ($a-\bar{c}$). For a firm, reducing the marginal cost c_i is among the most important tasks.

According to Cournot's conclusions, total output grows and the price of goods decreases when increasing the number of firms. In the linear model case, these conclusions are true; see formulas (2) and (3). For the nonlinear model, Cournot's conclusions remain valid under the conditions formulated in [44].

It follows from formula (3) that

$$P(Q^*)-\bar{c}\to 0\ \text{as}\ n\to\infty.$$

In other words, oligopoly turns into perfect competition with an unlimited increase in the number of firms. Oligopoly with firms producing a homogeneous good is called quasi-competition if total output grows and the price of goods decreases when increasing the number of firms. Under the identical cost functions of all firms, quasi-competitiveness alone is not sufficient for an oligopoly to become perfect competition with an unlimited increase in the number of firms; for details, see [45]. The case of different cost functions of firms was studied in [46]; as was shown therein, oligopoly is quasi-competitive under the conditions ensuring the uniqueness of Cournot equilibrium in the paper [16].

The dependence of the price of goods and the share of output produced by one firm on the number of firms was further examined in [47]. The author [48] also studied the transition from Cournot to perfect competition when the number of firms increases, including the stability of Cournot equilibrium; as it turned out, the system is stable if the capacities of firms decrease with increasing the number of firms. In the case of collusion between firms, the same issues were considered in [49] using conjectural variations, by analogy with [10]. The author [50] provided a rigorous mathematical proof of the so-called "folk theorem": if firms are small relative to the market, then there exists a Cournot equilibrium approximately representing perfect competition. By assumption, firms can enter and leave the market freely. Note that the average cost function of firms was adopted in the analysis. The research [50] was continued in [51], where the asymptotic properties of Cournot equilibria were analyzed. For the equilibrium with free entry of firms into the market, the dependence of the number of firms and total output on the demand and cost functions was investigated in [52].

The works [53–55] addressed society's welfare under Cournot competition. As it turned out within the mathematical models considered, the free entry of firms into the market can worsen rather than improve society's welfare. Consumer surplus and irrecoverable losses were also studied in [56]. With a certain simplification of the mathematical model, explicit elementary functions-based expressions were derived in [57] for some efficiency measures. In [58], firms producing several goods were considered to compare Cournot competition and collusion in terms of consumer surplus, total surplus, and the profit losses of firms. Note that uncertainty, which can make collusion inefficient, was neglected in that paper. The publication [59] was devoted to a Cournot oligopoly with free entry of new firms into the market and the following issues: how do the number of firms, the output of an individual firm, and total output depend on entry cost and on the market size? In [60], the restrictions of authorities for new firms entering the market were analyzed; one goal of the authorities is to improve society's welfare.

The paper [61] continued the studies initiated in [54, 55] that the entry of new firms into the market may be redundant. The R&D investment of firms to reduce costs was studied; the literature on Cournot equilibria was linked to the previous literature on R&D investment. In [62], models with firms choosing between two production technologies were considered. (The type of the cost function depends on this choice.) As was shown, in some cases, Cournot equilibria may not exist.

3. COURNOT AND BERTRAND OLIGOPOLIES

When studying Bertrand oligopoly, a usual assumption is that firms produce differentiated goods. In this case, firm i sets a price p_i for its good, $i=1,..., n$. The demand q_i for this good depends on all prices $p_1,..., p_n$. Then the profit of firm I is given by

$$\pi_i\left(p_1,..., p_n\right)=p_i q_i\left(p_1,..., p_n\right)-C_i\left(q_i\left(p_1,..., p_n\right)\right).$$

Firm i seeks to maximize the profit π_i. It is required to find a Nash equilibrium in the class of pure strategy profiles $p_1,..., p_n$.

The reason to consider firms with differentiated goods in Bertrand oligopoly is that, in the case of producing a homogeneous good, the unique Nash equilibrium may be the prices equal to marginal cost, i.e., zero profit for each firm. (Of course, to obtain this result, some assumptions must be accepted regarding the functions q_i and C_i; for example, see [2, 63, 64].) The conclusions concerning Bertrand oligopoly, particularly the comparison of Cournot and Bertrand oligopolies, essentially depend on whether the goods are substitutes or complements. (For the definition of substitutes and complements, we refer, e.g., to the monograph [1].)

Firms producing differentiated goods can also be considered within Cournot oligopoly. Then the profit of firm i is given by (instead of (1))

$$\pi_i\left(q_1,...,q_n\right)=P_i\left(q_1,...,q_n\right)q_i-C_i\left(q_i\right).$$

The equilibrium prices of goods under Cournot competition were compared with their counterparts under Bertrand competition, e.g., in the book [65, *pp. 68–78*]. In the duopoly problem considered therein, the equilibrium prices under Cournot competition are not lower than those under Bertrand competition. This result was confirmed in [66]; a linear duopoly model was investigated in detail, including the comparison of equilibrium prices and outputs, the profits of firms, consumer surplus, and total surplus. In the paper [67], similar results were established for nonlinear duopoly models. However, as was shown by the author [68], the conclusions of [66] may break when considering an arbitrary number of firms (not two).

In [69], Cournot and Bertrand equilibria were compared for firms producing a homogeneous good. Cournot and Bertrand oligopolies were compared by various criteria in [70–75] and other publications. In the linear case, Cournot competition and Bertrand competition were compared in detail in the paper [76]. Innovations to reduce cost were compared for Cournot and Bertrand oligopolies in [77]. The author [78] considered a duopoly with firms producing differentiated goods; the marginal cost of production was reduced using R&D expenditures both in a given firm and (with some factor) in the other firm; society's welfare under the two types of competition (Cournot and Bertrand) was investigated as a function of this reduction factor. Also, Cournot competition was compared with Bertrand competition in [79], but without using the concept of Nash equilibrium.

The paper [80] overviewed models in which output is the strategy of some firms and price is the strategy of the other firms (the so-called Cournot–Bertrand oligopoly). In addition, such an oligopoly was discussed in [81]. The paper [82] considered the environmental impact of production; the amount of pollution was compared for three models (Cournot duopoly, Bertrand duopoly, and Cournot–Bertrand duopoly).

4. STRATEGY CHOICE UNDER UNCERTAINTY. INFORMATION SHARING BETWEEN FIRMS. SELLING LICENSES

Consider a duopoly with $P(Q)=a-Q$ $=a-(q_1+q_2)$ and zero cost for both firms. Let a be a random variable taking values 60 and 120 with an equal probability of 0.5. The first firm is informed, i.e., knows the current (true) value a, and determines its output by formula (4) based on this value. The second (uninformed) firm proceeds from the fixed value $a=90$ and also determines its output by formula (4).

For the current value $a=120$, we obtain

$$q_1=40,\ q_2=30,\ Q=70.$$

Then the formula $P=a-Q$ yields $P=50$. The resulting profits are $\pi_1=Pq_1=2\,000$ and $\pi_2=Pq_2=1\,500$. But if the first firm informs the second one about the current value a, then $q_2=40$, $Q=80$, and $P=40$; in this case, $\pi_1=\pi_2=1\,600$. In other words, the profit of the first firm decreases whereas the profit of the second firm increases.

For the current value $a=60$, we obtain

$$q_1=20,\ q_2=30,\ Q=50.$$

Then the formula $P=a-Q$ yields $P=10$. The resulting profits are $\pi_1=Pq_1=200$ and $\pi_2=Pq_2=300$. But if the first firm informs the second one about the current value a, then $q_2=20$, $Q=40$, and $P=20$; in this case, $\pi_1=\pi_2=400$.

Thus, the average profit of the first firm is $0.5(2\,000+200)=1\,100$ if it does not share information with the second firm and is $0.5(1\,600+400)=1\,000$ otherwise. That is, the first firm benefits nothing from sharing information with the second firm. Similarly, for the second firm, the average profit is $0.5(1\,500+300)=900$ if it receives no information from the first firm and is $0.5(1\,600+400)=1\,000$ otherwise. Therefore, information sharing is beneficial for the second firm.

This example can be somewhat generalized. Suppose that the first and second firms have information ($a+\xi_1$) and ($a+\xi_2$) about the value a, respectively, where ξ_1 and ξ_2 are random variables with zero mean. Is it beneficial for the firms to share this information? Does the answer depend on the correlation of the random variables ξ_1 and ξ_2? The information may be more complex than just a single numerical parameter. Is it then beneficial to share some part of this information with another firm? Which part should it be?

For an arbitrary number of firms, such a problem with a random variable a was considered in [83]. Assume that the inverse demand function and cost functions of all firms are linear. The main question addressed in the paper is as follows: it is beneficial for the firms to research the market jointly? As it turned out, the benefit of reduced market research costs may be smaller than the losses due to the information available to competitors. In addition, the impact of market research on consumer surplus was studied. This impact is found to be positive, i.e., market research increases consumer satisfaction. A similar model for duopoly was considered in [84], and the authors arrived at the following conclusion: information sharing between firms increases the correlation of their outputs, reducing their expected profits. The difference between the works [85] and [84] is the use of Gaussian random variables. As a result, in some cases, the exact values of means can be calculated. Besides, information sharing between firms in [85] was associated with subsequent collusion to reduce output.

For a similar problem, the possibility of partial information sharing between firms was considered in [86]. The author [87] compared Cournot and Bertrand competition. Also, information sharing under Cournot and Bertrand competition was examined in [88]. In that paper, a duopoly with firms producing differentiated goods was considered in the following statement: the inverse demand functions are linear; the free terms in the price equations form a random vector with the 2D Gaussian distribution. In [89], similar problems were considered for the equation $P=a-bQ$ with b as a random variable. The example at the beginning of this section shows that the expected profit of a more informed firm is higher. In [90], an example of the opposite nature was provided for a duopoly with firms producing differentiated goods and uncertainty in some cross-effects.

Similar problems, but when exchanging information about cost rather than market demand, were considered in [91]. According to the conclusions, within the linear model, such information sharing increases the expected profits of firms but decreases the expected surplus of consumers. The author [92] studied problems where firms may (or may not) share both cost and market demand information with other firms. Theorems on the existence of Cournot equilibria and on convergence to perfect competition as the number of firms tends to infinity were obtained. The paper [93] challenged the conclusion of [91] that information sharing between firms reduces expected consumer surplus. In [94], the profits of firms depend on their strategies (outputs and prices) and, moreover, on the unknown state of the environment (each firm receives some signal about this state). The equilibria obtained with and without information sharing about these signals were compared.

The potentially disorienting character of information shared with other firms was discussed in [95]. The paper [96] presented many of the previous results on information sharing between firms within a unified model. In [97], similar issues were touched upon with respect to the Cournot–Bertrand duopoly (the mixed case with outputs and prices as the strategies of different firms). Several researchers considered the problem where each firm maximizes revenue or some weighted average of profit and revenue. In this case, the interests of owners and managers running the firm are separated. In [98], the reasonability of information sharing between firms in such a model was analyzed, including the effect of information dissemination on society's welfare.

In the problem described in [99], firms receive noisy signals about market demand and cost and determine Cournot equilibrium outputs based on these signals. Using the strong law of large numbers, it was proved that total output converges almost surely to the total output corresponding to perfect competition as the number of firms tends to infinity. Note that the linear model with the same marginal cost for all firms was adopted therein. The paper [100] developed the results and approaches from [92] and [99]; a continuum of firms was considered, and some convergence results were obtained for Bayesian–Nash equilibria.

A duopoly with firms producing differentiated goods was studied in [101]. In this model, the inverse demand function is subjected to random disturbances: the prices of goods under given outputs become random variables. The strategy of each firm takes the following form. First, the firm selects the type of strategy (the output or price of the good); then it decides on the output or price, respectively. All decisions are made simultaneously by both firms. Some results on the existence of Bayesian–Nash equilibria and on the expected profits of firms were given. A similar problem was considered in [102], where the firm's decision is related to the degree of substitutability of goods.

In [103], the following model was presented. The uncertainty regarding demand and cost is described by some probability space (Ω, F, μ). The awareness of firm i is given by a σ-subalgebra F_i included in the σ-algebra F. It was discussed under which conditions the better-informed firm (i.e., the firm with a larger σ-subalgebra F_i) would obtain a higher expected profit. In [104], the existence of Bayesian–Nash equilibria was investigated within the same problem statement.

The paper [105] considered a linear model with all completely known parameters. Each firm believes that all other firms will determine their outputs with probability p based on Cournot equilibrium and with probability $(1 - p)$ in another way, $0 < p < 1$. The pessimistic firm supposes "the worst-case" way whereas the optimist firm "the best-case" way. The best responses of firms as well as their possible outputs, prices, and profits were examined.

In [106], the inverse demand function has the form $P(Q) = \max(a - Q, 0)$, where a is a random variable taking two values ("large" and "small" with equal probabilities of 0.5). The ranges ensuring the existence of a unique Cournot equilibrium and exactly two Cournot equilibria were found. The expected profits and expected total surplus in such a model were compared with those in the model without uncertainty (i.e., in the model with the same "large" and "small" values). The paper [107] confirmed the conclusion of [106] that under uncertainty regarding market demand, the requirement of nonnegative prices may lead to a non-unique Cournot equilibrium. However, for models with free sales (firms can sell less goods than they produce), such non-uniqueness does not arise. In [108], uncertainty was related to the capacity of firms. In [109], Pareto equilibria in Cournot oligopoly were studied under uncertainty regarding demand.

In the publication [110], under a limited demand d for a good, the profit to be maximized is given by

$$\pi_i(q_i, Q) = P(Q)\min\left(q_i, \frac{q_i d}{Q}\right) - C_i(q_i),$$

where $C_i(q_i) = c_i q_i$ (c.f. formula (1)). Cournot equilibria were determined in the duopoly with this profit function. The demand d_t was then assumed to be a continuous-time random process satisfying some stochastic differential equation, $t \in [0, T]$. For random processes q_{1t} and q_{2t}, a control problem was studied to maximize functionals expressing the total profits of firms over the time interval from 0 to *T* with discounting.

The paper [111] was devoted to a duopoly where each firm produces two goods, *X* and *Y*. For technological reasons, the output of each good is a certain share of the firm's total output: $q_{iX} = \gamma q_i$ and $q_{iY} = (1-\gamma) q_i$, where $0 < \gamma < 1$, $i = 1, 2$. The inverse demand function is specific for each good:

$$P_X = a_X - b_X (q_{1X} + q_{2X}) \text{ and}$$

$$P_Y = a_Y - b_Y (q_{1Y} + q_{2Y}).$$

The random vector (a_X, a_Y) was supposed to obey the bivariate Gaussian distribution. Cournot competition was considered: firms determine the values q_1 and q_2. It was investigated how the availability of information about the true inverse demand function would affect the firm's expected profit.

In [112], a nonlinear duopoly model was considered as follows. One firm (e.g., firm 1) receives revenue from selling a license to another firm (firm 2). By doing so, firm 2 reduces cost. In this setup, formula (1) is replaced by

$$\pi_1(q_1, q_2) = P(Q) q_1 - C_1(q_1) + (r q_2 + f),$$

$$\pi_2(q_1, q_2) = P(Q) q_2 - C_1(q_2) - (r q_2 + f),$$

where f is a fixed fee and r denotes the royalty. It was studied under what conditions firm 2 would buy a license, both firms would remain active, or firm 1 would become a monopolist. The cases of drastic and non-drastic technologies were examined separately.

In [112], the linear model analysis from [113] (with calculations similar to those yielding (2)–(5)) was transferred to the nonlinear case. For the linear case, the author [113] presented the relative advantages and disadvantages of fixed fees and royalties. Also, his considerations were transferred to the nonlinear quadratic case in [114]; the differences from the linear case were discussed.

The researchers [115] studied a similar problem within the linear duopoly model with non-drastic technologies and foreign (firm 1) and domestic (firm 2) agents. Consumer surplus and society's welfare were studied for different cases ($f > 0$ and $r = 0$; $f = 0$ and $r > 0$; $f > 0$ and $r > 0$). The paper [116] investigated the possible impact of the limited capacity of the firm selling the license. In [117], the sales of license was considered in the case where firms 1 and 2 produce differentiated goods; Cournot competition and Bertrand competition were compared. Also, Cournot competition and Bertrand competition in the sales of licenses with R&D cost were compared in [118]; partially, the conclusions are opposite to those of [66].

5. CARTEL FORMATION

If a cartel is formed, the optimization problem will change. Cartelized firms seek to maximize total profit instead of their individual profits. The distribution of total profit among the firms may or may not be included in the analysis. Thus, within the mathematical model under consideration, cartel formation does not differ from the merger of firms, although it is not the same in reality.

Consider the model with the linear inverse demand function $P(Q) = a - bQ$ and the same cost function $C_i(q_i) = c q_i$ for all firms. By $\pi(m)$ we denote the profit of one firm given m firms in the market. Assume that *n* firms initially operate in the market, and k firms form a cartel. Total profit is distributed equally among all cartelized firms. Then the member firms benefit from forming the cartel if

$$\pi(n-k+1) - k\pi(n)$$

$$= \frac{(a-c)^2}{b} \left(\frac{1}{(n-k+2)^2} - \frac{k}{(n+1)^2} \right) > 0.$$

(Here, formula (5) is used.)

The paper [119] established a result called paradoxical in many subsequent works: the difference above is positive only for k close enough to n, approximately for $k \geq 0.8n$.

One drawback is that this model does not distinguish the cartel from other firms and neglects the "big" size of the newly created firm. To eliminate this drawback, the authors [120, 121] considered the same problem with the cost function $C_i(q_i) = \frac{q_i^2}{2k_i}$, where k_i is the capital of firm i. The impact of cartel formation on society's welfare was also studied in [121].

The problem of cartel formation under Bertrand competition was considered in [122]. The results obtained therein radically differ from those of [119]. Cartel formation increases the profits of all firms, while the profits of non-cartelized firms are greater than the profits of cartelized firms. In addition, an increase in the number of cartelized firms raises the profits of the cartelized firm.

Within the model proposed in [123], n firms produce a homogeneous good; among them, k firms are leaders and the rest are followers. Followers compete in the Cournot sense given the total output of leaders. Leaders realize this fact (i.e., they know the reaction functions of followers to their outputs) and also have Cournot competition among themselves by determining their outputs. Initially, all firms are assumed to be identical. The author [124] studied society's welfare for such a model. Similar problems with quadratic cost functions were considered in [125].

The following model was used to analyze cartel stability in [126]. In an economy sector, there are n identical firms producing a homogeneous good. Then k firms form a cartel and set the price of the good to maximize their total profit. Let $\pi_c(k)$ and $\pi_f(k)$ denote the profit of the cartelized and non-cartelized firms, respectively. A cartel is called *internally stable* if $\pi_c(k)/k \ge \pi_f(k-1)$, i.e., a cartelized firm will not increase its profit by leaving the cartel. A cartel is called *externally stable* if $\pi_f(k) \ge \pi_c(k+1)/(k+1)$, i.e., a non-cartelized firm will not increase its profit by joining the cartel. A cartel is called stable if it is both internally stable and externally stable. The paper [127] examined the process of forming a stable cartel by combining the ideas from [126, 128]. Note that neither Cournot nor Bertrand competition was considered in [126]. The concept of cartel stability in relation to Cournot competition, Bertrand competition, and leader-follower games was studied in many publications; for example, see [129–133].

Choosing a coalition structure that best matches the interests of players is a central problem in cooperative game theory. Such approaches are also applied to oligopoly games. In [134], a very general notion of equilibrium was considered; it includes the endogenous determination of the best coalition structure under a given set of admissible coalition structures. Particular cases of this equilibrium are Nash equilibrium for a noncooperative game and the core for a cooperative game with non-transferable utility. In both cases mentioned, a given set of admissible coalition structures is a singleton. In the former case, only coalitions consisting of one player are allowed. In the latter case, only coalitions consisting of all players are allowed. The paper [135] separately considered coalitions with positive and negative externalities, i.e., players not included in the coalition benefit (lose, respectively) from coalition formation. Stability was analyzed for different coalition formation rules (e.g., the consent of all coalition members is required for new entrants to join the coalition; the coalition can break up or merge with other coalitions, etc.). Research works focused on *partition function form games* (instead of cooperative games in the characteristic function form) were surveyed in [136].

6. PRACTICAL CALCULATION OF COURNOT EQUILIBRIA AND THEIR USE IN APPLICATIONS

Consider n firms producing m goods; q_i, the output of firm i, is an m-dimensional vector; L_i, the capacity of firm i, is also an m-dimensional vector; $P(Q)$ is the m-dimensional vector of prices; $P(Q)q_i$ in formula (1) is understood as an inner product. If $q^* = \left(q_1^*, \ldots, q_n^*\right)$ is an equilibrium, then q_i^* maximizes $\pi_i\left(q_1^*, \ldots, q_{i-1}^*, q_i, q_{i+1}^*, \ldots, q_n^*\right)$ as a function of the argument q_i for any i. A common approach is to find Cournot equilibrium by solving the complementarity problem.

Let π_i be a concave and continuously differentiable function of the argument q_i. Then for any $k = 1, \ldots, m$, the partial derivative $\frac{\partial \pi_i}{\partial q_{ik}}\left(q_i^*\right)$ is nonpositive if $q_{ik}^* = 0$; equal to zero if $0 < q_{ik}^* < L_{ik}$; and nonnegative if $q_{ik}^* = L_{ik}$. In other words, there exist values u_{ik} and v_{ik} such that

$$\frac{\partial \pi_i}{\partial q_{ik}}\left(q_i^*\right) + u_{ik} - v_{ik} = 0, \qquad (6)$$

where $u_{ik} \ge 0$ if $q_{ik}^* = 0$ and $u_{ik} = 0$ if $q_{ik}^* > 0$; $v_{ik} = 0$ if $q_{ik}^* < L_{ik}$ and $v_{ik} \ge 0$ if $q_{ik}^* = L_{ik}$. By the definition of u_{ik} and v_{ik}, it follows that

$$u_{ik} q_{ik}^* = 0 \text{ and } v_{ik}\left(L_{ik} - q_{ik}^*\right) = 0. \qquad (7)$$

With the m-dimensional vectors $u_i = \left(u_{i1}, \ldots, u_{im}\right)'$ and $v_i = \left(v_{i1}, \ldots, v_{im}\right)'$ (prime means transpose), the relation (6) can be written as

$$\operatorname{grad} \pi_i\left(q_i^*\right) + u_i - v_i = 0,$$

and the relations (7) as the zero scalar products

$$u_i q_i^* = 0 \text{ and } v_i\left(L_i - q_i^*\right) = 0.$$

In addition, $u_i \ge 0$, $v_i \ge 0$, $q_i^* \ge 0$, and $L_i - q_i^* \ge 0$. (For vectors, the nonstrict inequalities with "≥ 0" hold componentwise.)

Consider the $2m$-dimensional vectors

$$w_i = \begin{pmatrix} q_i \\ v_i \end{pmatrix}, \; i = 1, \ldots, n,$$

and the function

$$f_i(w_i) = \begin{pmatrix} v_i - \text{grad}\ \pi_i(q_i) \\ L_i - q_i \end{pmatrix}$$

that translates $2m$-dimensional vectors into $2m$-dimensional vectors. For $q_i = q_i^*$, the scalar product satisfies the equality

$$w_i f_i(w_i) = 0, \quad i = 1,..., n. \tag{8}$$

We introduce the notations

$$q = \begin{pmatrix} q_1 \\ \vdots \\ q_n \end{pmatrix}, \ v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \ L = \begin{pmatrix} L_1 \\ \vdots \\ L_n \end{pmatrix},$$

$$\text{grad}\ \pi(q) = \begin{pmatrix} \text{grad}\ \pi_1(q_1) \\ \vdots \\ \text{grad}\ \pi_n(q_n) \end{pmatrix}, \ w = \begin{pmatrix} q \\ v \end{pmatrix},$$

and $f(w) = \begin{pmatrix} v - \text{grad}\ \pi(q) \\ L - q \end{pmatrix}$.

Thus, Cournot equilibria can be determined by solving the complementarity problem: it is required to find vectors $w \geq 0$ such that $f(w) \geq 0$ and

$$w f(w) = 0.$$

The existence of solutions for complementarity problems was explored in [137, 138]. Due to (8), if q^* is a Cournot equilibrium, then there exists a vector v^* with nonnegative components such that the vector

$$w^* = \begin{pmatrix} q^* \\ v^* \end{pmatrix}$$

is the solution of the complementarity problem.

Some literature on the application of this approach for finding Cournot equilibria was cited, e.g., in [139]. Based on the same ideas, Cournot equilibria can also be determined in the case of nondifferentiable demand functions [140]. The authors [141, 142] proposed to calculate Cournot equilibria as solutions of some mathematical programming problems. In [143], another algorithm for finding Cournot equilibria was presented using the ideas from [137, 138, 142]. Based on complementarity problems, the paper [35] considered the existence and uniqueness of Cournot equilibria under price constraints as well as algorithms for finding Cournot equilibria.

Also, we note the following aspect: if firms produce a homogeneous good, then in some cases, Cournot equilibria can be determined by directly analyzing the multivalued mapping that associates with each strategy profile of other firms the best responses of a given firm [144]. Cournot equilibria can be found using the so-called tâtonnement process, which also involves best responses [145, *pp. 84–97*]. For firms with differentiated goods, this process was discussed in [146].

In [147], solutions of complementarity problems were adopted to study Cournot competition in electricity markets with uncertainty. Cournot equilibria were also used to analyze electricity markets in [139, 140, 148].

The researchers [149] compared Cournot competition and Bertrand competition in the software industry. One firm sells a platform (e.g., an operating system), and the other two firms supply application software. The comparison results for this problem quite differ from the ones obtained commonly.

In [150], competition between air carriers and railroad companies in high-speed transportation was treated as a Cournot duopoly. A model with linear inverse demand functions and zero cost was applied. By assumption, air carriers can have price discrimination (sell tickets to different groups of passengers at different prices) whereas railroad companies cannot. As it turned out, price discrimination increases the profit of air carriers. Consumer surplus and society's welfare were also studied.

CONCLUSIONS

This survey has presented the main sections of Cournot oligopoly as a mathematical discipline:

– the existence, uniqueness, and stability of equilibria;

– market entry for new firms and efficiency issues;

– a comparison of Cournot oligopoly, Bertrand oligopoly, and Cournot–Bertrand oligopoly;

– consideration of uncertainty by probabilistic methods (model parameters, deviations of real outputs from the planned ones, etc.);

– information sharing among firms and the sale of licenses;

– cartel formation;

– numerical methods for determining Cournot equilibria, primarily for multidimensional problems.

(Consideration of uncertainty by fuzzy set theory methods goes beyond the scope of this paper.)

The application of this mathematical theory in various branches of the economy has been considered.

The oligopoly theory has been evolving toward including more and more real processes in the mathematical model. In this way, recommendations for control and management can be developed. Mathematical modeling of information sharing among firms and cartel formation is important to determine which information exchange or the level of interaction between firms best matches the interests of firms. Also, such mathematical models are crucial from the society's welfare point of view to prevent a significant increase in prices for goods.

REFERENCES

1. Varian, H.R., *Intermediate Microeconomics. A Modern Approach*, 4th ed., New York: W.W. Norton, 1996.
2. Jehle, J.A. and Reny, P.J., *Advanced Microeconomic Theory*, 3rd ed., Upper Saddle River: Prentice Hall, 2011.
3. Dixit, A. and Stern, N., Oligopoly and Welfare: A Unified Presentation with Applications to Trade and Development, *European Economic Review*, 1982, vol. 19, pp. 123–143.
4. Theocharis, R.D., On the Stability of the Cournot Solution on the Oligopoly Problem, *Review of Economic Studies*, 1960, vol. 27, pp. 133–134.
5. Negishi, T., The Stability of a Competitive Economy: A Survey Article, *Econometrica*, 1962, vol. 30, pp. 635–669.
6. Fisher, F.M., The Stability of the Cournot Oligopoly Solution: The Effects of Speeds of Adjustment and Increasing Marginal Costs, *Review of Economic Studies*, 1961, vol. 28, pp. 125–135.
7. Hahn, F.H., The Stability of the Cournot Oligopoly Solution, *Review of Economic Studies*, 1962, vol. 29, pp. 329–333.
8. Okuguchi, K., The Stability of the Cournot Oligopoly Solution: A Further Generalization, *Review of Economic Studies*, 1964, vol. 31, pp. 143–146.
9. Hadar, J., Stability of Oligopoly with Product Differentiation, *Review of Economic Studies*, 1966, vol. 33, pp. 57–60.
10. Seade, J., The Stability of Cournot Revisited, *J. Econ. Theory*, 1980, vol. 23, pp. 15–27.
11. Chekanskii, A.N. and Frolova, N.L., *Mikroekonomika. Promezhutochnyi uroven'* (Microeconomics. An Intermediate Level), Moscow: Infra-M, 2005. (In Russian.)
12. Okuguchi, K. and Yamazaki, T., Global Stability of Unique Nash Equilibrium in Cournot Oligopoly and Public Good Game, *J. of Economic Dynamics and Control*, 2008, vol. 32, pp. 1204–1211.
13. Gao, B. and Du, Y., Exploring General Equilibrium Points for Cournot Model, *Discrete Dynamics in Nature and Society*, 2018, art. ID 7630395.
14. Soldatos, G., Marginal Utility and Cournot Stability under Market Entry, *Studies in Business and Economics*, 2021, vol. 16, pp. 187–197.
15. Frank, C.R. and Quandt, R.E., On the Existence of Cournot Equilibrium, *Int. Economic Review*, 1963, vol. 4, pp. 92–96.
16. Okuguchi, K. and Suzumura, K., Uniqueness of the Cournot Oligopoly Equilibrium: A Note, *Economic Studies Quarterly*, 1971, vol. 22, pp. 81–83.
17. Rosen, J.B., Existence and Uniqueness of Equilibrium Points for Concave N-person Games, *Econometrica*, 1965, vol. 33, pp. 520–534.
18. Szidarovszky, F. and Yakowitz, S., A New Proof of the Existence and Uniqueness of the Cournot Equilibrium, *Int. Economic Review*, 1977, vol. 18, pp. 787–789.
19. Okuguchi, K. and Szidarovszky, F., *The Theory of Oligopoly with Multi-Product Firms*, 2nd ed., Berlin: Springer, 1999.
20. Fujimoto, T., Cournot Oligopoly Model with Multi-product Firms, *Economic Studies Quarterly*, 1989, vol. 40, pp. 364–368.
21. Ewerhart, C., Cournot Games with Biconcave Demand, *Games and Economic Behavior*, 2014, vol. 85, pp. 37–47.
22. Roberts, J. and Sonnenschein, H., On the Existence of Cournot Equilibrium without Concave Profit Functions, *J. of Economic Theory*, 1976, vol. 13, pp. 112–117.
23. Novshek, W. and Sonnenschein, H., Cournot and Walras Equilibrium, *J. of Economic Theory*, 1978, vol. 19, pp. 223–266.
24. Radi, D., Walrasian versus Cournot Behavior in an Oligopoly of Boundedly Rational Firms, *J. of Evolutionary Economics*, 2017, vol. 27, pp. 933–961.
25. Novshek, W., On the Existence of Cournot Equilibria, *Review of Economic Studies*, 1985, vol. 52, pp. 85–98.
26. Bamón, R. and Frayssé, J., Existence of Cournot Equilibrium in Large Markets, *Econometrica*, 1985, vol. 53, pp. 587–597.
27. Gaudet, G. and Salant, S.W., Uniqueness of Cournot Equilibrium: New Results from Old Methods, *Review of Economic Studies*, 1991, vol. 58, pp. 399–404.
28. Corchón, L.C. and Torregrosa, R.J., Cournot Equilibrium Revisited, *Mathematical Social Sciences*, 2020, vol. 106, pp. 1–10.
29. Asplund, M., Risk-Averse Firms in Oligopoly, *Int. J. of Industrial Organization*, 2002, vol. 20, pp. 995–1012.
30. Loury, G., A Theory of 'Oil'igopoly: Cournot Equilibrium in Exhaustive Resource Markets with Fixed Supplies, *Int. Economic Review*, 1986, vol. 27, pp. 285–300.
31. Okuguchi, K., Labor-Managed Cournot Oligopoly with Product Differentiation, *J. of Economics*, 1992, vol. 56, pp. 197–208.
32. Okuguchi, K. and Yamazaki, T., Existence of Unique Equilibrium in Cournot Mixed Oligopoly, *Int. Game Theory Review*, 2018, vol. 20, art. no. 1750035.
33. Bergstrom, T.C. and Varian, H.R., Two Remarks on Cournot Equilibria, *Economics Lett.*, 1985, vol. 19, pp. 5–8.
34. Katz, M.L. and Rosen, H.S., Tax Analysis in an Oligopoly Model, *Public Finance Quarterly*, 1985, vol. 13(1), pp. 3–20.
35. Han, L. and Liu, A.L., On Nash-Cournot Games with Price Caps, *Operations Research Letters*, 2013, vol. 41, pp. 92–97.
36. Deo, S. and Corbett, C.J., Cournot Competition under Yield Uncertainty: The Case of the U.S. Influenza Vaccine Market, *Manufacturing and Service Operations Management*, 2009, vol. 11, pp. 563–576.
37. Liu, L., Correlated Equilibrium of Cournot Oligopoly Competition, *J. of Economic Theory*, 1996, vol. 68, pp. 544–548.
38. Yi, S.-S., On the Existence of a Unique Correlated Equilibrium in Cournot Oligopoly, *Economics Letters*, 1997, vol. 54, pp. 235–239.
39. Novikov, D.A. and Chkhartishvili, A.G., *Refleksivnye igry* (Reflexive Games), Moscow: SINTEG, 2003. (In Russian.)
40. Algazin, G.I. and Algazina, Y.G., Reflexive Dynamics in the Cournot Oligopoly under Uncertainty, *Automation and Remote Control*, 2020, vol. 81, no. 2, pp. 287–301.
41. Algazin, G.I. and Algazina, Y.G., Reflexion Processes and Equilibrium in an Oligopoly Model with a Leader, *Automation and Remote Control*, 2020, vol. 81, no. 7, pp. 1258–1270.
42. Geraskin, M.I., Analysis of Equilibria in a Nonlinear Oligopoly Model, *Automation and Remote Control*, 2022, vol. 83, no. 8, pp. 1261–1277.
43. Ougolnitsky, G.A. and Usov, A.B., The Interaction of Economic Agents in Cournot Duopoly Models under Ecological

Conditions: A Comparison of Organizational Modes, *Automation and Remote Control*, 2023, vol. 84, no. 2, pp. 175–189.
44. Frank, C.R., Entry in a Cournot Market, *Review of Economic Studies*, 1965, vol. 32, pp. 245–250.
45. Ruffin, R.J., Cournot Oligopoly and Competitive Behaviour, *Review of Economic Studies*, 1971, vol. 38, pp. 493–502.
46. Okuguchi, K., Quasi-competitiveness and Cournot Oligopoly, *Review of Economic Studies*, 1973, vol. 40, pp. 145–148.
47. Amir, R. and Lambson, V.E., On the Effects of Entry in Cournot Markets, *Review of Economic Studies*, 2000, vol. 67, pp. 235–254.
48. Puu, T., On the Stability of Cournot Equilibrium When the Number of Competitors Increases, *J. of Economic Behavior and Organization*, 2008, vol. 66, pp. 445–456.
49. Seade, J., On the Effects of Entry, *Econometrica*, 1980, vol. 48, pp. 479–489.
50. Novshek, W., Cournot Equilibria with Free Entry, *Review of Economic Studies*, 1980, vol. 47, pp. 473–486.
51. Ushio, Y., Cournot Equilibrium with Free Entry: The Case of Decreasing Average Cost Functions, *Review of Economic Studies*, 1983, vol. 50, pp. 347–354.
52. Corchón, L. and Fradera, I., Comparative Statics in Cournot Free Entry Equilibrium, *Math. Soc. Sci.*, 2002, vol. 44, no. 2, pp. 155–168.
53. von Weizsäcker, C.C., A Welfare Analysis of Barriers to Entry, *Bell J. of Economics*, 1980, vol. 11, pp. 399–420.
54. Mankiw, N.G. and Whinston, M.D., Free Entry and Social Efficiency of Entry, *Rand J. of Economics*, 1986, vol. 17, pp. 48–58.
55. Suzumura, K. and Kiyono, K., Entry Barriers and Economic Welfare, *Review of Economic Studies*, 1987, vol. 54, pp. 157–167.
56. Anderson, S.P. and Renault, R., Efficiency and Surplus Bounds in Cournot Competition, *J. Economic Theory*, 2003, vol. 113, pp. 253–264.
57. Corchón, L., Welfare Losses under Cournot Competition, *Int. J. Ind. Organ*., 2008, vol. 26, pp. 1120–1131.
58. Kluberg, J. and Perakis, G., Generalized Quantity Competition for Multiple Products and Loss of Efficiency, *Operations Research*, 2012, vol. 60, pp. 335–350.
59. Dastidar, K.G. and Marjit, S., Market Size, Entry Costs and Free Cournot Equilibrium, *J. of Economics*, 2022, vol. 136, pp. 97–114.
60. Goerke, L., Partisan Competition Authorities, Cournot-Oligopoly, and Endogenous Market Structure, *Southern Economic J*., 2022, vol. 89, pp. 238–270.
61. Okuno-Fujiwara, M. and Suzumura, K., Symmetric Cournot Oligopoly and Economic Welfare: A Synthesis, *Econ. Theory*, 1993, vol. 3, pp. 43–59.
62. Götz, G., Market Size, Technology Choice, and the Existence of Free-Entry Cournot Equilibrium, *J. of Institutional and Theoretical Economics*, 2005, vol. 161, pp. 503–521.
63. Avdasheva, S.B. and Rozanova, N.M., *Teoriya organizatsii otraslevykh rynkov* (Theory of Organization of Industrial Markets), Moscow: Magistr, 1998. (In Russian.)
64. Dastidar, K.G., On the Existence of Pure Strategy Bertrand Equilibrium, *Economic Theory*, 1995, vol. 5, pp. 19–32.
65. Shubik, M. and Levitan, R.E., *Market Structure and Behavior*, Cambridge: Harvard Univ. Press, 1980.
66. Singh, N. and Vives, X., Price and Quantity Competition in a Differentiated Duopoly, *Rand J. of Economics*, 1984, vol. 15, pp. 546–554.
67. Cheng, L., Comparing Bertrand and Cournot Equilibria: A Geometric Approach, *Rand J. of Economics*, 1985, vol. 16, pp. 146–152.
68. Häckner, J., A Note on Price and Quantity Competition in Differentiated Oligopolies, *J. of Economic Theory*, 2000, vol. 93, pp. 233–239.
69. Dastidar, K.G., Comparing Cournot and Bertrand in a Homogeneous Product Market, *J. Econ. Theory*, 1997, vol. 75, pp. 205–212.
70. Amir, R. and Jin, J.Y., Cournot and Bertrand Equilibria Compared: Substitutability, Complementarity and Concavity, *Int. J. of Industrial Organization*, 2001, vol. 19, pp. 303–317.
71. Hsu, J. and Wang, X.H., On Welfare under Cournot and Bertrand Competition in Differentiated Oligopolies, *Review of Industrial Organization*, 2005, vol. 27, pp. 185–191.
72. Zanchettin, P., Differentiated Duopoly with Asymmetric Costs, *J. of Economics & Management Strategy*, 2006, vol. 15, pp. 999–1015.
73. Correa-López, M., Price and Quantity Competition in a Differentiated Duopoly with Upstream Suppliers, *J. of Economics & Management Strategy*, 2007, vol. 16, pp. 469–505.
74. Farahat, A. and Perakis, G., A Comparison of Bertrand and Cournot Profits in Oligopolies with Differentiated Products, *Operations Research*, 2011, vol. 59, pp. 507–513.
75. Chang, M.C. and Peng, H.-P., Cournot and Bertrand Equilibria Compared: A Critical Review and an Extension from the Output-Structure Viewpoint, *Japanese Economic Review*, 2012, vol. 63, pp. 467–496.
76. Choné, P. and Linnemer, L., The Quasilinear Quadratic Utility Model: An Overview, *CESifo Working Paper no. 7640*, Munich: Munich Society for the Promotion of Economic Research, CESifo GmbH, 2019.
77. Bester, H. and Petrakis, E., The Incentives for Cost Reduction in a Differentiated Industry, *Int. J. of Industrial Organization*, 1993, vol. 11, pp. 519–534.
78. Qiu, L.D., On the Dynamic Efficiency of Bertrand and Cournot Equilibria, *J. of Economic Theory*, 1997, vol. 75(1), pp. 213–229.
79. Bulow, J.I., Geanakoplos, J.D., and Klemperer, P.D., Multimarket Oligopoly: Strategic Substitutes and Complements, *J. of Political Economy*, 1985, vol. 93, pp. 488–511.
80. Tremblay, C.H. and Tremblay, V.J., Oligopoly Games and the Cournot–Bertrand Model, *J. of Economic Surveys*, 2019, vol. 33, pp. 1555–1577.
81. Friedman, J.W., On the Strategic Importance of Prices versus Quantities, *Rand J. of Economics*, 1988, vol. 19, pp. 607–622.
82. Asproudis, E. and Filippiadis, E., Environmental Technological Choice in a Cournot–Bertrand Model, *J. of Industry, Competition & Trade*, 2021, vol. 21, pp. 43–58.
83. Ponssard, J.-P., The Strategic Role of Information on the Demand Function in an Oligopolistic Market, *Management Science*, 1979, vol. 25, pp. 243–250.
84. Novshek, W. and Sonnenschein, H., Fulfilled Expectations Cournot Duopoly with Information Acquisition and Release, *Bell J. of Economics*, 1982, vol. 13, pp. 214–218.
85. Clarke, R.N., Collusion and the Incentives for Information Sharing, *Bell J. of Economics*, 1983, vol. 14, pp. 383–394.
86. Gal-Or, E., Information Sharing in Oligopoly, *Econometrica*, 1985, vol. 53, pp. 329–343.
87. Gal-Or, E., Information Transmission – Cournot and Bertrand Equilibria, *Review of Economic Studies*, 1986, vol. 53, pp. 85–92.
88. Sakai, Y., Cournot and Bertrand Equilibria under Imperfect Information, *J. Econ.*, 1986, vol. 46, pp. 213–232.
89. Malueg, D.A. and Tsutsui, S.O., Duopoly Information Exchange: The Case of Unknown Slope, *Int. J. of Industrial Organization*, 1996, vol. 14, pp. 119–136.

90. Chokler, A., Hon-Snir, S., Kim, M., and Shitovitz, B., Information Disadvantage in Linear Cournot Duopolies with Differentiated Products, *Int. J. of Industrial Organization*, 2006, vol. 24, pp. 785–793.
91. Shapiro, C., Exchange of Cost Information in Oligopoly, *Review of Economic Studies*, 1986, vol. 53, pp. 433–446.
92. Li, L., Cournot Oligopoly with Information Sharing, *Rand J. of Economics*, 1985, vol. 16, pp. 521–536.
93. Sakai, Y. and Yamato, T., Oligopoly, Information and Welfare, *J. of Economics*, 1989, vol. 49, pp. 3–24.
94. Vives, X., Trade Association Disclosure Rules, Incentives to Share Information, and Welfare, *Rand J. of Economics*, 1990, vol. 21, pp. 409–430.
95. Ziv, A., Information Sharing in Oligopoly: The Truth-Telling Problem, *Rand J. of Economics*, 1993, vol. 24, pp. 455–465.
96. Raith, M. A General Model of Information Sharing in Oligopoly, *J. of Economic Theory*, 1996, vol. 71, pp. 260–288.
97. Kopel, M. and Putz, E.M., Information Sharing in a Cournot–Bertrand Duopoly, *Manage. Decis. Econ.*, 2021, vol. 42, pp. 1645–1655.
98. Theilen, B., Delegation and Information Sharing in Cournot Duopoly, *J. of Economics*, 2007, vol. 92, pp. 21–50.
99. Palfrey, T., Uncertainty Resolution, Private Information Aggregation, and the Cournot Competitive Limit, *Review of Economic Studies*, 1985, vol. 52, pp. 69–83.
100. Vives, X., Aggregation of Information in Large Cournot Markets, *Econometrica*, 1988, vol. 56, pp. 851–876.
101. Klemperer, P. and Meyer, M., Price Competition vs. Quantity Competition: The Role of Uncertainty, *Rand J. of Economics*, 1986, vol. 17, pp. 618–638.
102. Reisinger, M. and Ressner, L., The Choice of Prices versus Quantities under Uncertainty, *J. of Economics & Management Strategy*, 2009, vol. 18, pp. 1155–1177.
103. Einy, E., Moreno, D., and Shitovitz, B., Information Advantage in Cournot Oligopoly, *J. of Economic Theory*, 2002, vol. 106, pp. 151–160.
104. Einy, E., Haimanko, O., Moreno, D., and Shitovitz, B., On the Existence of Bayesian Cournot Equilibrium, *Games and Economic Behavior*, 2010, vol. 68, pp. 77–94.
105. Fontini, F., Cournot Oligopoly under Strategic Uncertainty with Optimistic and Pessimistic Firms, *Metroeconomica*, 2005, vol. 56, pp. 318–333.
106. Lagerlöf, J.N.M., Insisting on a Non-negative Price: Oligopoly, Uncertainty, Welfare, and Multiple Equilibria, *Int. J. Ind. Organ.*, 2007, vol. 25, pp. 861–875.
107. Grimm, V., Cournot Competition under Uncertainty, *Working Paper*, Köln: University of Cologne, 2008.
108. Richter, J., Incomplete Information in Cournot Oligopoly: The Case of Unknown Production Capacities, *EWI Working Paper no. 13/01*, Köln: Institute of Energy Economics, the University of Cologne (EWI), 2013.
109. Caraballo, M.A., Mármol, A.M., Monroy, L., and Buitrago, E.M., Cournot Competition under Uncertainty: Conservative and Optimistic Equilibria, *Rev. Econ. Design*, 2015, vol. 19, pp. 145–165.
110. Polak, I. and Privault, N., Cournot Games with Limited Demand: From Multiple Equilibria to Stochastic Equilibrium, *Appl. Math. Optim.*, 2020, vol. 81, pp. 195–220.
111. Okura, M., Cournot Competition in the Joint Products Market under Demand Uncertainty, *Manage. Decis. Econ.*, 2021, vol. 42, pp. 1105–1116.
112. Sen, D. and Stamatopoulos, G., Licensing under General Demand and Cost Functions, *European J. of Operational Research*, 2016, vol. 253, pp. 673–680.
113. Wang, X.H., Fee versus Royalty Licensing in a Differentiated Cournot Duopoly, *J. Econ. Bus*., 2002, vol. 54, pp. 253–266.
114. Fauli-Oller, R. and Sandonís, J., Fee versus Royalty Licensing in a Cournot Duopoly with Increasing Marginal Costs, *Manchester School*, 2022, vol. 90, pp. 439–452.
115. Kabiraj, A. and Kabiraj, T., Tariff Induced Licensing Contracts, Consumers' Surplus and Welfare, *Economic Modelling*, 2017, vol. 60, pp. 439–447.
116. Colombo, S., Filippini, L., and Sen, D., Patent Licensing and Capacity in a Cournot Model, *Review of Industrial Organization*, 2022, vol. 62, pp. 1–18.
117. Ferreira, F.A., Ferreira, F., and Bode, O.R., Licensing under Cournot vs. Bertrand Competition, *Economic Research – Ekonomska Istraživanja*, 2021, vol. 34, pp. 651–1675.
118. Li, C. and Ji, X., Innovation Licensing and Price vs. Quantity Competition, *Economic Modelling*, 2010, vol. 27, pp. 746–754.
119. Salant, S.W., Switzer, S., and Reynolds, R.J., Losses from Horizontal Merger: The Effects of an Exogenous Change in Industry Structure on Cournot–Nash Equilibrium, *Quarterly J. of Economics*, 1983, vol. 98, no. 2, pp. 185–199.
120. Perry, M.K. and Porter, R.H., Oligopoly and the Incentive for Horizontal Merger, *American Economic Review*, 1985, vol. 75, no. 1, pp. 219–227.
121. McAfee, R.P. and Williams, M.A., Horizontal Mergers and Antitrust Policy, *J. of Industrial Economics*, 1992, vol. 40, pp. 181–187.
122. Deneckere, R. and Davidson, C., Incentives to Form Coalitions with Bertrand Competition, *Rand J. of Economics*, 1985, vol. 16, no. 4, pp. 473–486.
123. Daughety, A.F., Beneficial Concentration, *American Economic Review*, 1990, vol. 80, no. 5, pp. 1231–1237.
124. Feltovich, N., Mergers, Welfare, and Concentration: Results from a Model of Stackelberg–Cournot Oligopoly, *Atlantic Economic J*., 2001, vol. 29, pp. 378–392.
125. Heywood, J.S. and McGinty, M., Leading and Merging: Convex Costs, Stackelberg, and the Merger Paradox, *Southern Economic J*., 2008, vol. 74, no. 3, pp. 879–893.
126. d'Aspremont, C., Jacquemin, A., Gabszewicz, J.J., and Weymark, J.A., On the Stability of Collusive Price Leadership, *Canadian J. of Economics*, 1983, vol. 16, no. 1, pp. 17–25.
127. Prokop, J., Process of Dominant Cartel Formation, *Int. J. of Industrial Organization*, 1999, vol. 17, pp. 241–257.
128. Selten, R., A Simple Model of Imperfect Competition, Where 4 Are Few and 6 Are Many, *Int. J. of Game Theory*, 1973, vol. 2, pp. 141–201.
129. Shaffer, S., Stable Cartels with a Cournot Fringe, *Southern J. of Economics*, 1995, vol. 61, pp. 744–754.
130. Thoron, S., Formation of Coalition-Proof Stable Cartels, *Canadian J. of Economics*, 1998, vol. 31, pp. 63–76.
131. Konishi, H. and Lin, P., Stable Cartels with a Cournot Fringe in a Symmetric Oligopoly, *Keio Economic Studies*, 1999, vol. 36, no. 2, pp. 1–10.
132. Escrihuela-Villar, M., A Note on Cartel Stability and Endogenous Sequencing with Tacit Collusion, *J. of Economics*, 2009, vol. 96, pp. 137–147.
133. Zu, L., Zhang, J., and Wang, S., The Size of Stable Cartels: An Analytical Approach, *Int. J. of Industrial Organization*, 2012, vol. 30, pp. 217–222.
134. Ichiishi, T., A Social Coalitional Equilibrium Existence Lemma, *Econometrica*, 1981, vol. 49, pp. 369–377.
135. Yi, S.-S., Stable Coalition Structures with Externalities, *Games and Economic Behavior*, 1997, vol. 20, pp. 201–237.

136. Abe, T., Cartel Formation in Cournot Competition with Asymmetric Costs: A Partition Function Approach, *Games*, 2021, vol. 12, art. no. 14.
137. Karamardian, S., The Nonlinear Complementarity Problem with Applications. Part 1, *J. of Optimization Theory and Applications*, 1969, vol. 4, pp. 87–98.
138. Karamardian, S., The Nonlinear Complementarity Problem with Applications. Part 2, *J. of Optimization Theory and Applications*, 1969, vol. 4, pp. 167–181.
139. Hobbs, B., Linear Complementarity Models of Nash–Cournot Competition in Bilateral and POOLCO Power Markets, *IEEE Transactions on Power Systems*, 2001, vol. 16, pp. 194–202.
140. Hobbs, B. and Pang, J., Nash–Cournot Equilibria in Electric Power Markets with Piecewise Linear Demand Functions and Joint Constraints, *Operations Research*, 2007, vol. 55, pp. 113–127.
141. Maňas, M., A Linear Oligopoly Game, *Econometrica*, 1972, vol. 40, pp. 917–922.
142. Murphy, F.H., Sherall, H.D., and Soyster, A.L., A Mathematical Programming Approach for Determining Oligopolistic Market Equilibrium, *Mathematical Programming*, 1982, vol. 24, pp. 92–106.
143. Harker, P.T., A Variational Inequality Approach for the Determination of Oligopolistic Market Equilibrium, *Mathematical Programming*, 1984, vol. 30, pp. 105–111.
144. Novshek, W., Finding All N-Firm Cournot Equilibria, *Int. Economic Review*, 1984, vol. 25, pp. 61–70.
145. Moulin, H., *Theorie des jeux pour l'economie et la politique*, Paris: Hermann, 1981. (In French.)
146. Zhang, A. and Zhang, Y., Stability of a Cournot–Nash Equilibrium: The Multiproduct Case, *J. of Mathematical Economics*, 1996, vol. 26, pp. 441–462.
147. Siriruk, P., Cournot Competition under Uncertainty in Power Markets, *PhD Dissertation*, Alabama: Auburn University, 2009.
148. Valenzuela, J. and Mazumdar, M., Cournot Prices Considering Generator Availability and Demand Uncertainty, *IEEE Transactions on Power Systems*, 2007, vol. 22, pp. 116–125.
149. Fanti, L. and Buccella, D., Cournot and Bertrand Competition in the Software Industry, *Economics Research International*, 2017, vol. 2017, art. ID 9752520.
150. Li, J., Li, X., and Ma, N., A Price Discrimination Based Cournot Game Model for High-Speed Rail and Airline, *J. of Intelligent and Fuzzy Systems*, 2021, vol. 41, pp. 4793–4801.

This paper was recommended for publication by RAS Academician D.A. Novikov, a member of the Editorial Board.

Received March 14, 2023,
and revised September 28, 2023.
Accepted September 28, 2023.

Author information

Shvedov, Alexey Sergeevich. Dr. Sci. (Phys.–Math.), National Research University Higher School of Economics, Moscow, Russia
✉ ashvedov@hse.ru
ORCID iD: https://orcid.org/0000-0002-8988-0432

Cite this paper

Shvedov, A.S., Cournot Oligopoly: Strategy Choice under Uncertainty and Other Problems. *Control Sciences* **5,** 18–31 (2023). http://doi.org/10.25728/cs.2023.5.2

Translated into English by *Alexander Yu. Mazurov*,
Cand. Sci. (Phys.–Math.),
Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
✉ alexander.mazurov08@gmail.com

DOI: http://doi.org/10.25728/cs.2023.5.3

CONTROL SYSTEM DESIGN FOR MOVING OBJECTS WITH CHANNEL SWITCHING

E.Yu. Abdullina and V.N. Efanov

Ufa University of Science and Technology, Ufa, Russia

✉ elzik86@mail.ru, ✉ efanov@mail.ru

Abstract. This paper considers the problem of designing control systems for moving objects with control channel switching. A generalized characteristic is proposed to eliminate jumps and impulses that may occur at switching instants. This characteristic describes the entire ensemble of system trajectories under control channel switching at an arbitrary random instant. A numerical inversion method is developed for the Laplace transform based on expanding the delta function into a series of exponential polynomials. With this method, the generalized characteristic of the system can be approximated by a given time domain. The exponential series description of the generalized system characteristic allows reducing the original design problem to a system of algebraic inequalities. A particular example of designing an automatic pitch control system for an aircraft with a normal overload limit is presented; as is shown, the entire ensemble of trajectories obtained for different channel switching instants belongs to a given time domain.

Keywords: design, switched systems, selector, moving object, trajectory, angle of attack.

INTRODUCTION

Control systems of moving objects have a number of peculiarities [1, 2]. One peculiarity is their multi-mode operation. In this case, the transition to the next operating mode is accompanied by a structural change in the control part of the system. In particular, such a situation arises in the programmed control of a moving object when it is necessary to restrict the maximum allowed values of the motion parameters, e.g., the limiting angle of attack in aircraft attitude control. During operation the system structure changes according to an accepted logic of channel switching; therefore, such systems are referred to as switched systems [3–10]. This is a class of multi-mode dynamic systems consisting of a family of continuous subsystems and a device that controls mode switching.

The permanently increasing interest in switched systems is due to their wide use in applications (control of electric power systems, aircraft, industrial processes, and many other areas, up to the development of intelligent components of control systems [11–17]). Also, there are several interesting phenomena occurring in such systems. Examples show that the stability of all individual modes is not sufficient for the stability of a system with arbitrary switching [18–21]. In other words, the stability of switched systems depends on the dynamics of the system in each mode and the consistency of the modes during their switching. In this regard, research efforts are concentrated on the stability and stabilizability of switched systems [22–26], as well as methods for designing controllers with guaranteed stability and control performance [27–31]. There exist two approaches to the design of such systems. In the first case, dynamic processes are specified by a system of differential equations, which form a finite ensemble of typical trajectories. The optimal system trajectory is constructed of separate sections of these typical trajectories by a switching device (a finite automaton with memory). The memory of the switching device stores the time intervals on which the motion will follow the selected trajectory and the sequence in which the selected trajectories will be switched. In the studies of such switched systems, the main attention is paid to the development of logical operation rules for the automaton to ensure a consistent change in the system state at the switching instant without jumps and bursts of the controlled variable.

The second approach to the design of switched systems is used when the switching rules are described by constraints on time and state, or are due to external impacts applied to the controlled object. One example is the need to restrict the maximum allowed values of the motion parameters; see the discussion above. In this case, structural switching can be considered a disturbance applied to the system. Therefore, the continuous subsystems of the system's control part should compensate such disturbances by matching the system trajectories when changing its operating modes. A distinctive feature is that such systems change their properties in a jump-like manner at unknown random instants. As a result, they are treated as systems with random structure [32–34]. Consequently, there is an entire ensemble of optimal trajectories, each corresponding to a particular switching instant. In this case, the averaged or generalized characteristics of the system are used to assess its dynamics [35–37]. However, the scope of these results is often restricted to analysis problems since the system characteristics in all structural states are supposed to be known. In this regard, we propose a generalized characteristic of a system with channel switching for design purposes: it allows deriving an analytical dependence between the parameters of the control part and the ensemble of system trajectories under channel switching at an arbitrary random instant.

This paper develops a control design algorithm to match the trajectories of systems with control channel switching using a generalized characteristic that describes the entire ensemble of their trajectories under control channel switching at an arbitrary random instant.

1. ANALYSIS OF SYSTEM DYNAMICS WITH CHANNEL SWITCHING AT RANDOM INSTANTS

Assume that in the control mode of the object's parameter $y(t)$, the system's trajectory is described by the equation

$$P_1(D)\,y(t)=Q_1(D)\,y_{\text{con}}(t), \qquad (1)$$

where $y_{\text{con}}(t)$ is the control program for the given parameter; $P_1(D)=a_n^{(1)}\dfrac{d^n}{dt^n}+\dots+a_1^{(1)}\dfrac{d}{dt}+a_0^{(1)}$ and $Q_1(D)=b_m^{(1)}\dfrac{d^m}{dt^m}+\dots+b_1^{(1)}\dfrac{d}{dt}+b_0^{(1)}$, where $a_i^{(1)}\left(i=\overline{0,n}\right)$, $b_j^{(1)}\left(j=\overline{0,m}\right)$ are the model parameters.

When switching to the limiting channel at some instant τ, the system trajectory with respect to the parameter $y(t)$ is described by the equation

$$P_2(D)\,y(t,\ \tau)=Q_2(D)\,y_{\lim}(t) \qquad (2)$$

with initial conditions determining the coincident system states at the control mode change instant:

$$y^{(r)}(t,\ \tau)\big|_{t=\tau}=y^{(r)}(t)\big|_{t=\tau};\ \left(r=0,1,\dots,n-1\right),$$

where $y_{\lim}(t)$ is the limiting program for the given parameter; $P_2(D)=a_n^{(2)}\dfrac{d^n}{dt^n}+\dots+a_1^{(2)}\dfrac{d}{dt}+a_0^{(2)}$ and $Q_2(D)=b_m^{(2)}\dfrac{d^m}{dt^m}+\dots+b_1^{(2)}\dfrac{d}{dt}+b_0^{(2)}$.

To match the system trajectories when switching the control channel to the limiting channel at an arbitrary random instant $\tau\in[0,+\infty)$, we consider the following generalized characteristic of the system:

$$E\left(y(t)\right)=\int_0^\infty y(t,\ \tau)f(\tau)d\tau, \qquad (3)$$

where $f(\tau)$ is the distribution function of the random instant τ.

On the time interval $0\le t\le\tau$, the system trajectory satisfies equation (1). In this case, $y(t,\ \tau)=y(t)$. On the time interval $0\le\tau\le t$, the limiting loop comes into operation and the system trajectory $y(t,\ \tau)$ will satisfy equation (2).

Assume that the control structure is switched with a constant intensity λ. Then the distribution function of the switching instant obeys the law $f(\tau)=\lambda e^{-\lambda\tau}$. We divide the integral (3) into two terms corresponding to the operating modes specified above:

$$E\left(y(t)\right)=\int_0^t y(t,\ \tau)f(\tau)d\tau+\int_t^\infty y(t)f(\tau)d\tau$$

$$=\int_0^t y(t,\ \tau)f(\tau)d\tau+y(t)e^{-\lambda t}.$$

To calculate the latter integral, we represent the solution of system (2) as the sum $y(t,\ \tau)=y^{(\text{I})}(t)+y^{(\text{II})}(t,\ \tau)$ of the partial solution of the inhomogeneous equation (2) with zero initial conditions and the general solution of the corresponding homogeneous equation with nonzero initial conditions. Since the initial conditions of system (2) remain valid for $y(t,\ \tau)$, it follows that

$$\left[y^{(\text{II})}(t,\ \tau)\right]^{(i)}\Big|_{t=\tau}=\left[y(t)\right]^{(i)}\Big|_{t=\tau}-y^{(I)}\left[(t)\right]^{(i)}\Big|_{t=\tau},$$

$$i=\overline{0,(n-1)}.$$

The solution of the homogeneous equation has the form [38]

$$y^{(\text{II})}(t,\ \tau)=\sum_{i=0}^{n-1}\sum_{j=0}^{n-i-1}Y^{(i)}(\tau)a_{i+j+1}^{(2)}\cdot\left[w_2(t-\tau)\right]^{(j)},$$

where $Y^{(i)}(\tau)=[y(\tau)-y^{(1)}(\tau)]^{(i)}$ and $w_2(t)=L^{-1}\left\{\dfrac{1}{P_2(s)}\right\}$.

As a result,

$$E(y(t))=y(t)e^{-\lambda t}+y^{(1)}(t)\left(1-e^{-\lambda t}\right)$$
$$+\sum_{i=0}^{n-1}\sum_{j=0}^{n-i-1}a_{i+j+1}^{(2)}\int_0^t Y^{(i)}(\tau)\lambda e^{-\lambda\tau}\cdot\left[w_2(t-\tau)\right]^{(j)}d\tau. \quad (4)$$

The generalized characteristic of the system similar to (4) was used in the papers [39, 40]. However, the desired system trajectories under channel switching were ensured therein by localizing the roots of the denominator of the image of this generalized characteristic in a given domain of the complex plane. Generally speaking, such an approach does not eliminate all jumps and bursts at the channel switching instant since the temporal characteristics depend on the image denominator and also on its numerator. In this regard, we require that the generalized characteristic $E(y(t))$ of the system lies within given limits:

$$E_1(t)\le E(y(t))\le E_2(t). \quad (5)$$

2. MATCHING CONTROL OF THE SYSTEM WITH CHANNEL SWITCHING

To satisfy condition (5), let us find the Laplace image for $E(y(t))$:

$$E(s)=L\{E(y(t))\}=L\left\{y(t)e^{-\lambda t}\right\}$$
$$+L\left\{y^{(1)}(t)\left(1-e^{-\lambda t}\right)\right\}$$
$$+L\left\{\sum_{i=0}^{n-1}\sum_{j=0}^{n-i-1}a_{i+j+1}^{(2)}\int_0^t Y^{(i)}(\tau)\lambda e^{-\lambda\tau}\left[w_2(t-\tau)\right]^{(j)}d\tau\right\}.$$

Consequently,

$$E(s)=y(s+\lambda)+\left(y^{(1)}(s)-y^{(1)}(s+\lambda)\right)$$
$$+\frac{\lambda[y(s+\lambda)-y^{(1)}(s+\lambda)]}{P_2(s)}\sum_{i=0}^{n-1}\sum_{j=0}^{n-i-1}a_{i+j+1}^{(2)}(s+\lambda)^i s^j.$$

The double sum in this expression can be transformed as follows:

$$\sum_{i=0}^{n-1}\sum_{j=0}^{n-i-1}a_{i+j+1}^{(2)}(s+\lambda)^i s^j=\sum_{k=1}^{n}\sum_{l=0}^{k-1}a_k^{(2)}(s+\lambda)^l s^{k-l-1}$$
$$=\sum_{k=1}^{n}a_k^{(2)}\frac{(s+\lambda)^k-s^k}{(s+\lambda)-s}=\frac{P_2(s+\lambda)-P_2(s)}{\lambda}.$$

Considering this relation, the desired image becomes

$$E(s)=y^{(1)}(s)+\left(y(s+\lambda)-y^{(1)}(s+\lambda)\right)\frac{P_2(s+\lambda)}{P_2(s)}.$$

The inverse Laplace transform should be performed to find the original function of the generalized characteristic $E(y(t))$ based on the image $E(s)$. However, it seems impossible to do in a general form because the image depends on the unknown parameters of the system's control part. Therefore, we employ a special numerical inversion method for the Laplace transform. Within this method, the delta function $\delta(t,\tau)$ is approximated by a partial sum of the series $\delta_q(t,\tau)=\sum_{i=1}^{q}d_i(t)\,\varphi_i(\tau)$, where

$$\varphi_i(t)=\sum_{j=1}^{i}c_{ij}\exp\left(-\beta(j-1)t\right),\ (\beta>0;\ i=1,2,\ldots),$$

is the set of orthonormal exponential polynomials with the weight $g(t)=\exp(-\alpha t)$, $\alpha\ge 0$.

According to the paper [41], the coefficients c_{ij} are given by

$$c_{i+1,j+1}=\frac{(-1)^{i+j}\,\Gamma(i+j+\delta+1)\sqrt{(\delta+2i+1)\beta}}{j!(i-j)!\Gamma(j+\delta+1)},$$
$$i,\ j=0,1,\ldots,$$

where $\Gamma(x)$ denotes the gamma function and $\delta=(\alpha-\beta)/\beta$.

The exponential polynomials $\varphi_i(t)$ can be obtained from any classical orthogonal polynomials $p_i(z)$ (Legendre, Laguerre, Hermite polynomials, etc. [42]) by the change of variable $z=\exp(-\beta t)$. With this change, the system trajectories are described over the entire time horizon $t\in[0,+\infty)$.

The partial-sum sequence $\delta_q(t,\tau)$ converges to the delta function $\delta(t,\tau)$. To demonstrate this fact, we introduce the auxiliary function $v(t)=g(t)E(y(t))$ for which

$$\lim_{q\to\infty}\int_0^\infty v(\tau)\delta_q(t,\tau)d\tau$$
$$=\lim_{q\to\infty}\sum_{i=1}^{q}g(t)\varphi_i(t)\int_0^\infty g(\tau)E(y(\tau))\varphi_i(\tau)d\tau.$$

The integral on the right-hand side is the formula for calculating the coefficients of the orthogonal series when expanding the function $E(y(t))$ with respect to the system of exponential polynomials

$$e_i[E]=\int_0^\infty g(\tau)E(y(\tau))\varphi_i(\tau)d\tau.$$

Consequently,

$$\lim_{q\to\infty}\int_0^\infty v(\tau)\delta_q(t,\tau)d\tau=\lim_{q\to\infty}g(t)\sum_{i=1}^q e_i(E)\varphi_i(t).$$

Since the orthogonal series is convergent for any square integrable function $E(y(t))$ with the weight $g(t)$, we obtain

$$\lim_{q\to\infty}\sum_{i=1}^q e_i(E)\varphi_i(t)=E(y(t))$$

and $\lim_{q\to\infty}\int_0^\infty v(\tau)\delta_q(t,\tau)d\tau=g(t)E(y(t))=v(t).$

Hence,

$$\lim_{q\to\infty}\int_0^\infty v(\tau)\delta_q(t,\tau)d\tau=\int_0^\infty v(\tau)\delta(t,\tau)d\tau=v(t),$$

and it follows that

$$\delta(t,\tau)=\lim_{q\to\infty}\sum_{i=1}^q g(t)\varphi_i(t)\varphi_i(\tau).$$

This expansion of the delta function can serve to invert the Laplace transform

$$E(s)=\int_0^\infty e^{-s\tau}E(y(\tau))d\tau.$$

For this purpose, we use the following transformations:

$$\int_0^\infty e^{-s\tau}E(y(\tau))\sum_{i=1}^q g(t)\varphi_i(t)\varphi_i(\tau)d\tau$$
$$=\sum_{i=1}^q g(t)\varphi_i(t)\int_0^\infty e^{-s\tau}E(y(\tau))\varphi_i(\tau)d\tau$$
$$=\sum_{i=1}^q g(t)\varphi_i(t)\int_0^\infty e^{-s\tau}E(y(\tau))\sum_{j=1}^i c_{ij}\exp(-\beta(j-1)t)d\tau$$
$$=\sum_{i=1}^q g(t)\varphi_i(t)\sum_{j=1}^i c_{ij}E(s+(j-1)\beta).$$

On the other hand,

$$\lim_{q\to\infty}\int_0^\infty e^{-s\tau}E(y(\tau))\sum_{i=1}^q g(t)\varphi_i(t)\varphi_i(\tau)d\tau$$
$$=\int_0^\infty e^{-s\tau}E(y(\tau))\delta_q(t,\tau)d\tau=e^{-st}E(y(t)).$$

Letting $s=\alpha$ gives

$$E(y(t))=\lim_{q\to\infty}\sum_{i=1}^q\varphi_i(t)\sum_{j=1}^i c_{ij}E(\alpha+(j-1)\beta)$$
$$=\lim_{q\to\infty}\sum_{i=1}^q e_i(E)\varphi_i(t). \quad (6)$$

Consequently, the coefficients of the series expansion of the function $E(y(t))$ with respect to the system of exponential polynomials are calculated based on the values of its image $E(s)$ at real-axis points:

$$e_i(E)=\sum_{j=1}^i c_{ij}E(\alpha+(j-1)\beta),\ i=1,2,\ldots$$

The generalized characteristic $E(y(t))$ written as the series expansion with respect to the system of exponential polynomials can be used to formalize the design problem of matching control in the system with channel switching.

3. CONTROL SYSTEM DESIGN WITH CHANNEL SWITCHING IN THE DESIRED DOMAIN OF TEMPORAL CHARACTERISTICS

We represent the partial sum of the exponential series (6) as

$$E(y(t))=\sum_{i=1}^q e_i(E)\varphi_i(t)=\sum_{k=1}^q r_k(E)\exp(-\beta(k-1)t),$$

where $r_k(E)=\sum_{i=1}^q\lambda_{ik}e_i(E)$.

Recommendations on choosing q (the number of series terms) to achieve the required approximation accuracy can be found in several sources (e.g., see the papers [42, 43]). Let us expand the boundaries of the desired domain of temporal characteristics into similar exponential series:

$$E_1(t)=\sum_{k=1}^q r_k(E_1)\exp(-\beta(k-1)t),$$
$$E_2(t)=\sum_{k=1}^q r_k(E_2)\exp(-\beta(k-1)t).$$

Then the system of inequalities (5) can be written as

$$E(y(t))-E_1(t)=\sum_{k=1}^q R_k^{(\mathrm{I})}(E)\exp(-\beta(k-1)t)\ge 0,$$
$$E_2(t)-E(y(t))=\sum_{k=1}^q R_k^{(\mathrm{II})}(E)\exp(-\beta(k-1)t)\ge 0,$$

where $R_k^{(\mathrm{I})}(E)=[r_k(E)-r_k(E_1)]$ and $R_k^{(\mathrm{II})}(E)=[r_k(E_2)-r_k(E)]$.

With the change of variable $z=\exp(-\beta t)$, the system of constraints takes the form

$$P_1(z)=\sum_{k=1}^q R_k^{(\mathrm{I})}(E)z^{(k-1)}\ge 0,\ P_2(z)=\sum_{k=1}^q R_k^{(\mathrm{II})}(E)z^{(k-1)}\ge 0.$$

For the polynomials $P_1(z)$ and $P_2(z)$ to be non-negative on the interval $[0, 1]$, it suffices to require the following: the polynomials must take a positive value at least at one point of this interval, and, in addition, all their real roots must lie to the right of the point $z = 1$.

According to Newton's theorem on the bounds of polynomial roots, the number $z = 1$ is a lower bound for the positive roots of the polynomials $P_1(z)$ and $P_2(z)$ if

$$\left[z^{q-1} P_1\left(\frac{1}{z}\right)\right]_{z=1}^{(p)} \geq 0, \left[z^{q-1} P_2\left(\frac{1}{z}\right)\right]_{z=1}^{(p)} \geq 0, \quad p = 0, 1, ..., q-1. \tag{7}$$

We require the polynomials $P_1(z)$ and $P_2(z)$ to be positive at $z = 0$, i.e., $P_1(0) = r_1(E) - r_1(E_1) > 0$, $P_2(0) = r_1(E_2) - r_1(E) > 0$.

Together with inequalities (7), these conditions lead to the set of constraints determining the belonging of the generalized characteristic $E(y(t))$ to the given domain:

$$\begin{gathered} r_1(E) - r_1(E_1) > 0, r_1(E_2) - r_1(E) > 0; \\ \sum_{i=1}^{q-p}\left[\frac{(q-i)!}{(q-p-i)!}\right]\left(r_i(E) - r_i(E_1)\right) \geq 0, \\ \sum_{i=1}^{q-p}\left[\frac{(q-i)!}{(q-p-i)!}\right]\left(r_i(E_2) - r_i(E)\right) \geq 0, \\ p = 0, 1, \ldots, q-2. \end{gathered} \tag{8}$$

Heuristic zero-order search algorithms (e.g., the Hooke–Jeeves pattern search method) are recommended to solve this system of algebraic inequalities.

To illustrate the proposed approach, we construct a pitch control system for an aircraft with a normal overload limit.

4. AUTOMATIC PITCH CONTROL SYSTEM DESIGN FOR AN AIRCRAFT WITH A NORMAL OVERLOAD LIMIT

Consider the pitch control system for an aircraft with the normal overload-limiting channel originally proposed in the paper [44]. The structural diagram of this system is shown in Fig. 1.

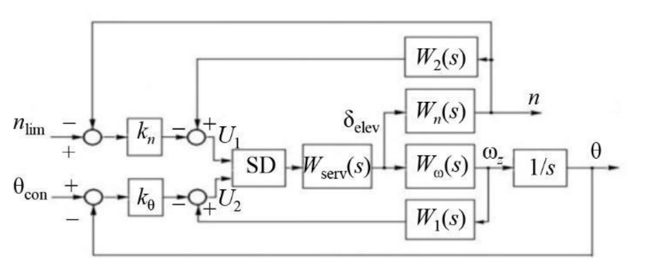

Fig. 1. The pitch control system with the normal overload-limiting channel.

The switching device (SD) supplies the maximum absolute value of the control signal to the servo drive:

$$U = \begin{cases} U_1 \text{ if } |U_1| > |U_2| \\ U_2 \text{ if } |U_2| > |U_1|. \end{cases}$$

The transfer functions of the aircraft to control the elevator δ_{elev} have the following form:

– for the pitch velocity ω_z, $W_\omega(s) = \dfrac{-(s+2.012)}{s^2 + 4.107s + 25.256}$;

– for the normal overload n, $W_n(s) = \dfrac{-1}{s^2 + 4.107s + 25.256}$.

The transfer function of the elevator servo drive is $W_{servo}(s) = \dfrac{10}{s}$.

Selecting a physically implementable astatic pitch autopilot with velocity feedback, we obtain the transfer function

$$W_1(s) = \frac{k_2 s^2 + k_1 s + k_0}{T_2 s^2 + T_1 s + 1},$$

where k_0, k_1, k_2, and k_θ are the autopilot gains; see Fig. 1.

In turn, we choose the transfer function $W_2(s)$ for the normal overload-limiting automaton in the form

$$W_2(s) = \frac{k_4 s^2 + k_3 s}{T_4 s^2 + T_3 s + 1},$$

where k_3, k_4, and k_n are the automaton's gains; see Fig. 1.

Using the introduced characteristics, we find the transfer function of the pitch control loop:

$$\Phi_1(s) = \frac{B_1(s)}{A_1(s)} = \frac{b_3^{(1)} s^3 + b_2^{(1)} s^2 + b_1^{(1)} s + b_0^{(1)}}{a_6^{(1)} s^6 + a_5^{(1)} s^5 + a_4^{(1)} s^4 + a_3^{(1)} s^3 + a_2^{(1)} s^2 + a_1^{(1)} s + a_0^{(1)}},$$

where

$$b_3^{(1)} = 10k_\theta T_2,\ b_2^{(1)} = 10k_\theta\left(2.012T_2 + T_1\right),$$

$$b_1^{(1)} = 10k_\theta\left(2.012T_1 + 1\right),\ b_0^{(1)} = 20.12k_\theta,\ a_6^{(1)} = T_2,$$

$$a_5^{(1)} = 4.107T_2 + T_1,$$

$$a_4^{(1)} = 25.256T_2 + 4.107T_1 + 10k_2 + 1,$$

$$a_3^{(1)} = 25.256T_1 + 10k_1 + 20.12k_2 + 10k_\theta T_2 + 4.107,$$

$$a_2^{(1)} = 10k_0 + 20.12k_1 + 20.12k_\theta T_2 + 10k_\theta T_1 + 25.256,$$

$$a_1^{(1)} = 20.12k_0 + 20.12k_\theta T_1 + 10k_\theta,\ \text{and}\ a_0^{(1)} = 20.12k_\theta.$$

Similarly, for the normal overload-limiting loop, we find the transfer function

$$\Phi_2(s) = \frac{B_2(s)}{A_2(s)} = \frac{b_3^{(2)}s^3 + b_2^{(2)}s^2 + b_1^{(2)}s + b_0^{(2)}}{a_6^{(2)}s^6 + a_5^{(2)}s^5 + a_4^{(2)}s^4 + a_3^{(2)}s^3 + a_2^{(2)}s^2 + a_1^{(2)}s},$$

where

$$b_3^{(2)} = 10k_nT_4,\ b_2^{(2)} = 10k_n\left(T_3 + 2.012T_4\right),$$

$$b_1^{(2)} = 10k_n\left(2.102T_3 + 1\right),\ b_0^{(2)} = 10k_n,\ a_6^{(2)} = T_4,$$

$$a_5^{(2)} = 4.107T_4 + T_3,\ a_4^{(2)} = 25.256T_4 + 4.107T_3 + 1,$$

$$a_3^{(2)} = 25.256T_3 + 10k_nT_4 + 10k_4 + 4.107,$$

$$a_2^{(2)} = 10k_3 + 10\,k_nT_3 + 25.256,\ \text{and}\ a_1^{(2)} = 10k_n.$$

Let the control channel be switched with the constant intensity $\lambda = 1\ \text{s}^{-1}$. We require that the generalized characteristic $E(y(t))$ of the system lies in the domain bounded by the functions $E_1(t) = 0.9\left(1 - 2\exp(-0.5t) + \exp(-t)\right)$ and $E_2(t) = 1.1\left(1 - \exp(-4t)\right)$. For the boundaries of this domain to belong to the exponential series basis, we select its parameters as follows: $\alpha = \beta = 0.5$ and $q = 9$.

The exponential series for the boundaries of the domains have the following coefficients:

$r_1(E_1) = 0.9$, $r_2(E_1) = -1.8$, $r_3(E_1) = 0.9$, $r_4(E_1) = 0$,

$r_5(E_1) = 0$, $r_6(E_1) = 0$, $r_7(E_1) = 0$, $r_8(E_1) = 0$, $r_9(E_1) = 0$,

$r_1(E_2) = 1.1$, $r_2(E_2) = 0$, $r_3(E_2) = 0$, $r_4(E_2) = 0$,

$r_5(E_2) = 0$, $r_6(E_2) = 0$, $r_7(E_2) = 0$, $r_8(E_2) = 0$,

and $r_9(E_2) = -1.1$.

In turn, for the generalized characteristic $E(y(t))$, these coefficients are calculated as

$$r_k(E) = \sum_{i=k}^{q}\sum_{j=1}^{i} c_{ij}c_{ik}E(0.5j),\ k = 1, 2, \dots, 9,$$

where

$$E(0.5j) = \frac{0.5jQ_1(0.5j+1)P_2(0.5j+1)}{0.5j(0.5j+1)P_1(0.5j+1)P_2(0.5j)} + \frac{\left[(0.5j+1)Q_2(0.5j) - 0.5jQ_2(0.5j)\right]P_1(0.5j+1)}{0.5j(0.5j+1)P_1(0.5j+1)P_2(0.5j)}.$$

Substituting these expressions into the system of inequalities (8) yields the set of constraints on the parameters of the control part. Solving the system of inequalities, we find the following parameter values:

$k_\theta = 60.49$, $k_n = 50.42$, $k_0 = 17.76$, $k_1 = 11.11$,

$k_2 = 1.11$, $k_3 = 16.49$, $k_4 = 1.99$, $T_1 = 0.50$,

$T_2 = 0.0005$, $T_3 = 0.002$, and $T_4 = 0.000001$.

Figure 2 shows the model of the control system designed in Matlab\Simulink.

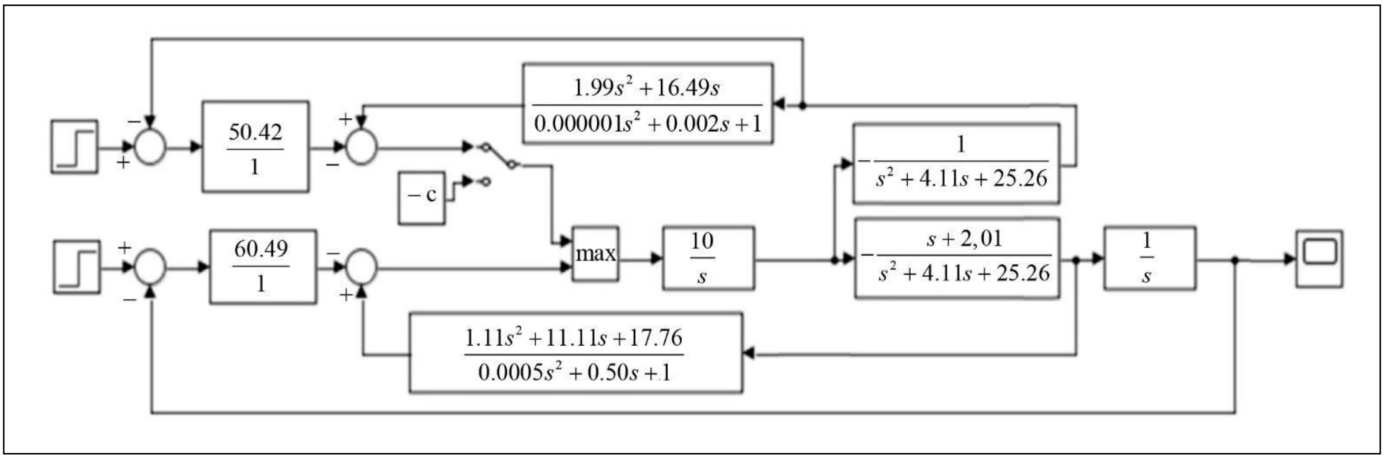

Fig. 2. The model of the control system.

The simulation results for this control system are demonstrated in Fig. 3.

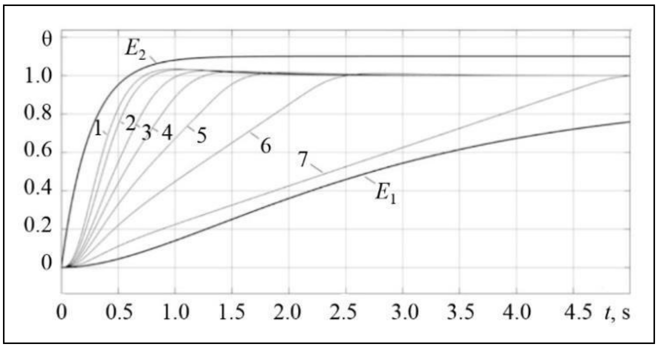

Fig. 3. Transients in the pitch control system with the normal overload-limiting channel.

The graphs in Fig. 3 correspond to the following switching instants: $(1 - \tau) = 0.1$ s, $(2 - \tau) = 0.3$ s, $(3 - \tau) = 0.6$ s, $(4 - \tau) = 0.8$ s, $(5 - \tau) = 1.3$ s, $(6 - \tau) = 2.1$ s, and $(7 - \tau) = 4.6$ s.

According to the simulation results, the system transients for different channel switching instants belong to the specified domain and keep the aperiodic

CONCLUSIONS

With the approach presented in this paper, the processes occurring in the control systems of moving objects with channel switching can be studied from a common point of view. For this purpose, we have proposed a generalized characteristic describing the entire ensemble of all output responses of the control system under all possible structural change instants (switching instants) of its control part. This characteristic can be used to design control systems for moving objects owing to a special numerical inversion method of the Laplace transform. As one example, an automatic pitch control system designed for an aircraft with a normal overload limit has illustrated the effectiveness of this approach.

REFERENCES

1. Lebedev, G.N., Nartov, B.K., and Chukanov, S.N., *Operativnyi kontrol' i upravlenie podvizhnymi ob"ektami* (Operational Control of Moving Objects), Moscow: Nauchtekhlitizdat, 2003. (In Russian.)
2. Pshihopov, V.H. and Medvedev, M.Yu., Multi-loop Adaptive Control of Mobile Objects in Solving Trajectory Tracking Tasks, *Control Sciences*, 2018, no. 6, pp. 62–72. (In Russian.)
3. Vasil'ev, S.N. and Malikov, A.I., About Some Results on the Stability of Switched and Hybrid Systems, in *Aktual'nye problemy mekhaniki sploshnoi sredy* (Topical Problems of Continuum Mechanics), Kazan: Foliant, 2011, vol. 1, pp. 23–81. (In Russian.)
4. Bortakovskii, A.S., Necessary Optimality Conditions for Switched Systems, *Journal of Computer and Systems Sciences International*, 2016, vol. 55, pp. 712–724. DOI: https://doi.org/10.1134/S1064230716050051.
5. Sun, Y., Zhang, C., Lu, X.-L., et al., Dynamic Optimization of Differential-algebraic Equations with Inequality Path Constraints, *Acta Automatica Sinica*, 2019, vol. 45, no. 5, pp. 897–905. DOI: 10.16383/j.aas.c180302.
6. Schwarz, D.E. and Lamour, R., A Projector Based Decoupling of DAEs Obtained from the Derivative Array, in *Progress in Differential-Algebraic Equations II*, Reis, T., Grundel, S., and Schöps, S., Eds., Cham: Springer, 2020, pp. 3–38. DOI: https://doi.org/10.1007/978-3-030-53905-4_1.
7. Demir, A. and Hanay, M.S., Numerical Analysis of Multidomain Systems: Coupled Nonlinear PDEs and DAEs with Noise, *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2018, vol. 37, no. 7, pp. 1445–1458. DOI: 10.1109/TCAD.2017.2753699.
8. Ñañez, P., Sanfelice, R.G., and Quijano, N., Notions and a Passivity Tool for Switched DAE Systems, *Proceedings of 2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, Melbourne, 2017, pp. 3612–3617. DOI: 10.1109/CDC.2017.8264190.
9. Yang, M., Lian, J., and Han, Y., Exponentially Passive Analysis of Switched Linear Systems with a Novel Storage Function, *Proceedings of 2018 Chinese Automation Congress (CAC)*, Xi'an, 2018, pp. 4014–4019. DOI: 10.1109/CAC.2018.8623256.
10. Chesi, G. and Colaneri, P., Structured Feedback Synthesis for Stability and Performance of Switched Systems, *IEEE Transactions on Automatic Control*, 2020, vol. 65, no. 11, pp. 4695–4709. DOI: 10.1109/TAC.2019.2962218.
11. Pappalardo, C.M. and Guida, D., On the Computational Methods for Solving the Differential-Algebraic Equations of Motion of Multibody Systems, *Machines*, 2018, vol. 6, no. 20, pp. 1–15.
12. Gai, W., Sun, C., Zhou, Y., and Zhang, J., A New Control Allocation Method Based on the Improved Grey Wolf Optimizer Algorithm for Aircraft with Multiple Actuators, *Proceedings of 2019 CAA Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS)*, Xiamen, 2019, pp. 438–442. DOI: 10.1109/SAFEPROCESS45799.2019.9213444.
13. Zhao, Y., Zhao, J., Fu, J., et al., Rate Bumpless Transfer Control for Switched Linear Systems with Stability and Its Application to Aero-Engine Control Design, *IEEE Transactions on Industrial Electronics*, 2020, vol. 67, no. 6, pp. 4900–4910. DOI: 10.1109/TIE.2019.2931222.
14. Li, J., Wei, G., Ding, D., and Li, Y., Quantized Control for Networked Switched Systems with a More General Switching Rule, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2020, vol. 50, no. 5, pp. 1909–1917. DOI: 10.1109/TSMC.2018.2791614.

15. Papadopoulos, A.V., Terraneo, F., Leva, A., and Prandini, M., Switched Control for Quantized Feedback Systems: Invariance and Limit Cycle Analysis, *IEEE Transactions on Automatic Control*, 2018, vol. 63, no. 11, pp. 3775–3786. DOI: 10.1109/TAC.2018.2797246.
16. Kuppusamy, S. and Joo, Y.H., Nonfragile Retarded Sampled-Data Switched Control of T–S Fuzzy Systems and Its Applications, *IEEE Transactions on Fuzzy Systems*, 2020, vol. 28, no. 10, pp. 2523–2532. DOI: 10.1109/TFUZZ.2019.2940432.
17. Fei, Z., Shi, S., Wang, T., and Ahn, C.K., Improved Stability Criteria for Discrete-Time Switched T–S Fuzzy Systems, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2021, vol. 51, no. 2, pp. 712–720. DOI: 10.1109/TSMC.2018.2882630.
18. Zhao, X., Yin, Y., Liu, L., and Sun, X., Stability Analysis and Delay Control for Switched Positive Linear Systems, *IEEE Transactions on Automatic Control*, 2018, vol. 63, no. 7, pp. 2184–2190. DOI: 10.1109/TAC.2017.2757460.
19. Wu, J., Yang, X., Zhang, C., and Li, J., Adaptive Finite-Time Control Design for a Class of Uncertain Nonlinearly Parameterized Switched Systems, *IEEE Access*, 2019, vol. 7, pp. 95941–95949. DOI: 10.1109/ACCESS.2019.2929841.
20. Zhu, Y. and Zheng, W.X., Multiple Lyapunov Functions Analysis Approach for Discrete-Time-Switched Piecewise-Affine Systems Under Dwell-Time Constraints, *IEEE Transactions on Automatic Control*, 2020, vol. 65, no. 5, pp. 2177–2184. DOI: 10.1109/TAC.2019.2938302.
21. Xiao, X., Zhou, L., Ho, D.W.C., and Lu, G., Event-Triggered Control of Continuous-Time Switched Linear Systems, *IEEE Transactions on Automatic Control*, 2019, vol. 64, no. 4, pp. 1710–1717. DOI: 10.1109/TAC.2018.2853569.
22. Trenn, S., Stabilization of Switched DAEs via Fast Switching, *PAMM: Proceedings in Applied Mathematics and Mechanics*, 2016, vol. 16, pp. 827–828. DOI: https://doi.org/10.1002/pamm.201610402.
23. Komaee, A., Stabilization of Linear Systems by Pulse-Width Modulation of Switching Actuators, *IEEE Transactions on Automatic Control*, 2020, vol. 65, no. 5, pp. 1969–1984. DOI: 10.1109/TAC.2019.2926943.
24. Wang, P. and Zhao, J., Feedback Dissipativity and Stabilization for Switched Positive Systems with a Combined Switching Law, *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2020, vol. 67, no. 11, pp. 2572–2576. DOI: 10.1109/TCSII.2019.2962283.
25. Wang, Z.-M., Wei, A., Zhao, X., et al., Stability Analysis of Discrete-Time Switched Systems with Unstable Modes: An Improved Ratio-Based Tradeoff Approach, *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2021, vol. 68, no. 1, pp. 431–435. DOI: 10.1109/TCSII.2020.3004400.
26. Harivanam, P.R. and Debasattam, P., Lie-Algebraic Criteria for Stability of Switched Systems of Differential Algebraic Equations (DAEs), *IEEE Control Systems Letters*, 2021, vol. 5(4), pp. 1333–1338. DOI: 10.1109/LCSYS.2020.3036577.
27. Bortakovskii, A.S., Necessary Optimality Conditions for Switching Systems, *Trudy Instituta Matematiki i Mekhaniki UrO RAN*, 2021, vol. 27, no. 2, pp. 67–78. (In Russian.)
28. Chen, Y. and Respondek, W., Geometric Analysis of Differential-Algebraic Equations via Linear Control Theory, *SIAM Journal on Control and Optimization*, 2021, vol. 59, no. 1, pp. 103–130. DOI: 10.1137/20M1329330.
29. Berger, T., Controlled Invariance for Nonlinear Differential–Algebraic Systems, *Automatica*, 2016, vol. 64, pp. 226–233.
30. Ilchmann, A., Leben, L., Witschel, J., and Worthmann, K., Optimal Control of Differential-Algebraic Equations from an Ordinary Differential Equation Perspective, *Optimal Control Applications and Methods*, 2019, vol. 40, no. 10, pp. 351–366. DOI: 10.1002/oca.2481.
31. Terasaki, S. and Kazuhiro, S., Minimal Controllability Problems on Linear Structural Descriptor Systems, *IEEE Transactions on Automatic Control*, 2021, vol. 67, no. 5, pp. 2522–2528. DOI: 10.1109/TAC.2021.3079359.
32. Bukhalev, V.A., Skrynnikov, A.A., and Boldinov, V.A., *Sistemy so sluchainoi skachkoobraznoi strukturoi* (Systems with Random Jump Structure), Moscow: Zhukovsky Academy, 2022. (In Russian.)
33. Averina, T.A. and Rybakov, K.A., Statistical Filtering Algorithms for Systems with Random Structure, *Russian Universities Reports. Mathematics*, 2020, vol. 25, no. 130, pp. 109–122. (In Russian.)
34. Bukhalev, V.A., Skrynnikov, A.A., and Boldinov, V.A., *Igrovoe upravlenie sistemami so sluchainoi skachkoobraznoi strukturoi* (Game-Theoretic Control of Systems with Random Jump Structure), Moscow: Fizmatlit, 2021. (In Russian.)
35. Sklyarevich, A.N. and Sklyarevich, F.A., *Veroyatnostnye modeli ob"ektov s vozmozhnymi izmeneniyami* (Probabilistic Models of Objects with Possible Changes), Riga: Zinatne, 1989. (In Russian.)
36. Fei, Z., Guan, C., and Zhao, X., Event-Triggered Dynamic Output Feedback Control for Switched Systems with Frequent Asynchronism, *IEEE Transactions on Automatic Control*, 2020, vol. 65, no. 7, pp. 3120–3127. DOI: 10.1109/TAC.2019.2945279.
37. Mostacciuolo, E., Vasca, F., and Baccari, S., Differential Algebraic Equations and Averaged Models for Switched Capacitor Converters with State Jumps, *IEEE Transactions on Power Electronics*, 2018, vol. 33, no. 4, pp. 3472–3483. DOI: 10.1109/TPEL.2017.2702389.
38. Doetsch, G. and Herschel, R., *Anleitung zum praktischen Gebrauch der Laplace-Transformation und der Z-Transformation*, München: R. Oldenbourg Verlag, 1967.
39. Abdullina, E.Yu. and Efanov, V.N., Roll Control of a Highly Maneuverable Aircraft under Conditions of Structural Uncertainty, *Journal of Instrument Engineering*, 2020, vol. 63, no. 1, pp. 26–34. (In Russian.)
40. Abdullina, E.Y. and Efanov, V.N., Synthesis of Pitch Angle Control System with Angle of Attack Limiting Channel, *Russian Aeronautics*, 2020, vol. 63, no. 1, pp. 25–32. DOI: 10.3103/S1068799820010043.
41. Denisenko, D.A. and Efanov, V.N., Synthesis of Robust Control Systems in the Environment of Orthogonal Functions of Exponential Type, Information and Control Systems, 2012, no. 4, pp. 52–58. (In Russian.)
42. Suetin, P.K., *Klassicheskie ortogonal'nye mnogochleny* (Classical Orthogonal Polynomials), Moscow: Fizmatlit, 2005. (In Russian.)
43. Szegő, G., *Orthogonal Polynomials*, American Mathematical Society Colloquium Publications, vol. 23, 4th ed., Providence: Amer. Math. Soc., 1975.
44. Petunin, V.I., Neugodnikova, L.M., and Abdullina, E.Yu., RF Patent 2014129734/11, *Byull. Izobret.*, 2015, no. 8.

This paper was recommended for publication by V.B. Pavlov, a member of the Editorial Board.

Received April 13, 2023,
and revised August 17, 2023.
Accepted August 30, 2023.

Author information

Abdullina, El'za Yunirovna. Postgraduate, Ufa University of Science and Technology, Ufa, Russia
✉ elzik86@mail.ru
ORCID iD: https://orcid.org/0009-0006-1800-2663

Efanov, Vladimir Nikolaevich. Dr. Sci. (Eng.), Ufa University of Science and Technology, Ufa, Russia
✉ efanov@mail.ru
ORCID iD: https://orcid.org/0000-0002-5917-2910

Cite this paper
Abdullina, E.Yu., Efanov, V.N., Control System Design for Moving Objects with Channel Switching. *Control Sciences* **5,** 32–40 (2023). http://doi.org/10.25728/cs.2023.5.3

Translated into English by *Alexander Yu. Mazurov*,
Cand. Sci. (Phys.–Math.),
Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
✉ alexander.mazurov08@gmail.com

DOI: http://doi.org/10.25728/cs.2023.5.4

STRATEGIC CAPABILITIES AS A DRIVER OF COMPETITIVENESS: A COMPARISON OF RUSSIAN AND GLOBAL COMPANIES

N.I. Guseva and O.Yu. Trubnikova

National Research University Higher School of Economics, Moscow, Russia

✉ natguseva@hse.ru, ✉ otrubnikova@hse.ru

Abstract. Competitiveness is a key measure of success for countries, industries, and companies. This study is devoted to the strategic capabilities underlying the competitiveness of companies in the Russian market. The theoretical foundations of strategic capabilities are conceptualized, and their role in the development of competitiveness through strengthening the competitive advantages of companies is emphasized. The research methodology of strategic capabilities is based on the leading publications of foreign and domestic scholars over the last 30 years. The empirical part of this study considers both the general and specific strategic capabilities of 30 Russian and multinational companies operating in the Russian market. According to the empirical results, Russian domestic companies focus on the development of one particular strategic capability only; Russian exporters have a more differentiated approach, developing a significant number of strategic capabilities simultaneously; finally, multinational companies endeavor to develop a balanced set of strategic capabilities. The most important strategic capabilities of companies in the Russian market are identified as follows: employee professionalism, quality control, innovations and unique technologies, corporate brand and reputation, market and trend understanding, customer orientation, and flexibility. The research results have practical value and can be used by Russian companies as key factors for increasing their competitiveness in the Russian market as well as in new foreign markets.

Keywords: company competitiveness, strategic capabilities, competitive advantage, multinational companies, Russian companies.

INTRODUCTION

High turbulence of the external environment and business competition create radically new conditions for Russian and multinational companies operating in the Russian market. This is especially true for the companies under serious sanction pressure on the Russian economy, which exacerbates the problem of managing the strategic capabilities of Russian companies. A company can gain a stronger competitive position by determining its strategic capabilities, designing them in accordance with the conditions of the external environment, and integrating them into a sustainable strategy [1]. With this view, strategic capabilities are considered a driver of competitiveness, enabling companies (in particular, Russian companies) to determine their "growth point."

In this paper, based on the results of a large-scale empirical study, we investigate the strategic capabilities of companies with different levels of internationalization operating in the Russian market as a driver of their competitiveness.

The analysis of 30 Russian and multinational companies operating in the Russian market reveals the dominant categories and subcategories of strategic capabilities in their competitiveness. We identify a set of strategic capabilities predominately used by the companies when building their competitiveness in the Russian market as well as analyze the key differences in the formation of competitiveness of the companies considering the level of their internationalization.

The structure of this study reflects the main theoretical (methodological) issues and the empirical results. The theoretical part of the study focuses on the basic conceptual aspects of strategic capabilities and approaches to their research. The methodology of the empirical part describes the object under study and the stages of analysis. The main results of this work are grouped according to the research questions indicated above. The key findings are presented in the Conclusions.

1. THEORETICAL FOUNDATIONS OF INVESTIGATING STRATEGIC CAPABILITIES

1.1. The Basic Conceptual Aspects of Strategic Capabilities

The concept of strategic capabilities rests on the following structural elements: necessary resources and *core competencies*, which should be unique and difficult to replicate [2]; *ordinary capabilities* and *dynamic capabilities* [3]; information and *organizational knowledge*, which is becoming a main asset of many organizations [4]; the organization's interaction with the external environment, which can provide the resources, competencies, and capabilities that the organization needs to achieve a particular strategic goal [5–7]. These elements, taken to a level of excellence beyond anything the company does, and especially better than any of its competitors [8], become strategic capabilities that enable an organization to survive and thrive. Building and renewing capabilities is at the heart of *competitive advantage* and organizational value creation [9]. V. Kumar traced the relationship between the external environment, strategy, and strategic capabilities that lead to superior organizational performance and sustainable competitive advantages [1].

Sustainable competitive advantages require the development of sustainable strategic capabilities, generally through a significant investment of time and resources from the company [8, 10, 11]. The sustainability of strategic capabilities consists of three components: *complexity*, *culture and history*, and *causal ambiguity*. Complexity is explained by the company's *internal* and *external linkages*, which can create obstacles for competitors to build capabilities or make them difficult to imitate [12]. Core competencies can become part of an organization's culture, i.e., competencies evolve over time in the so-called *path dependency* [13]; thus, they become specific to the organization and cannot be imitated. Finally, competitors find it difficult to identify the causes and outcomes underlying the organization's performance and its advantages because of *characteristic ambiguity* or *linkage ambiguity* within and outside the organization [14].

Strategic capabilities are directly related to the company's strategy, and their emergence has led to the development of such concepts as *strategic intent* (an expression of organization's intentions, plans, and ways to achieve them) and *strategic fit* (the extent to which a company matches its capabilities and resources with the opportunities available in the external environment) [15].

No capability can be the only basis for sustainable competitiveness. A. Johne characterized strategic capabilities as a "complex set" that allows companies to carry out their activities [16]. Thus, the set of strategic capabilities should be considered a driver of the competitiveness of companies through forming new and strengthening current competitive advantages.

1.2. Approaches to Investigating Strategic Capabilities

Strategic capabilities are a complex multilevel concept and manifest themselves in various forms [17–20]. Therefore, diverse approaches are needed to assess the construct of strategic capabilities. The following points of view define the differences in these approaches: the functional area in which strategic capabilities are applied; the capability hierarchy; and the focal unit of analysis.

Most authors adopt the functional typology of capabilities to study specific functional areas and places of their application within an organization. In particular, this conclusion is evident from a significant pool of empirical studies of dynamic and organizational capabilities [21]. Strategic capabilities are also divided according to the functional criterion, but there is no common approach among the authors to distinguish a uniform set of functions (Table 1). Most of the functional areas separated out are identical or overlapping. Organizations may try to access the necessary resources and capabilities beyond their borders or seek geographical expansion; hence, the purposeful management of organizational internationalization becomes significant, also indicating the relationship of strategic capabilities with the external environment and their dynamic component.

Another aspect of strategic capabilities that has received much attention is their location in the capability hierarchy. Consideration of the company's capability architecture assists in understanding the nature of strategic capabilities [26, 27]. Most of the models reflecting the capability hierarchy are three-level and include resources, operational and functional capabilities, core competencies, and dynamic capabilities. However, these models do not explicitly show the place of strategic capabilities in the architecture (Table 2).

Table 1

Approaches to assessing strategic capabilities depending on the functional area

Author(s)	Categories of strategic capabilities
Lenz, 1980 [5]	Post-sales service, retail sales, physical distribution, manufacturing, design, research and development
Day, 1994 [18]	*Outside-in* processes: market sensing, customer linking, channel bonding, technology monitoring *Spanning* processes: customer order fulfillment, pricing, purchasing, customer service delivery, new product/service development, strategy development *Inside-out* processes: financial management, cost control, technology development, integrated logistics, manufacturing/transformation processes, human resources management, environment health and safety
Hafeez, Zhang, and Malak, 2002 [22]	Design, purchasing, manufacturing, sales and marketing, R&D, finance, management, etc.
Inan & Bititci, 2015 [23]	R&D capability, innovation capability, product development capability, environmental scanning capability, networking capability, alliancing and acquisition capability, imitation/replication capability, reconfiguration capability, knowledge development/learning capability, marketing capability
Van Looy et al, 2012 [24]	Modeling, management, deployment, optimisation, culture, structure
Cuervo-Cazurra et al, 2020 [7]	Obtaining resources, product/service capabilities, operations and management, marketing, managing the external environment
LeanIX, 2022 [25]	Strategic management, customer relationships, product & service development, production, procurement & logistics, enterprise support, marketing & sales, finance & controlling

Table 2

Approaches to determining the capability hierarchy

Author(s)	Complexity →			
	Level 1	Level 2	Level 3	Level 4
Collis, 1994 [28]	First category capabilities – Functional capabilities (static)	Second category capabilities – Capabilities related to dynamics (dynamic)	Third category capabilities – Creative capabilities	Meta-capabilities – The ability to develop the capability to understand industry contexts better
Teece et al., 1997 [9]	Ordinary capabilities	Dynamic capabilities	–	–
Dannels, 2002 [29]	Competencies	Second level capabilities	Dynamic abilities	–
Winter, 2003 [30]	Competencies	Second-order capabilities	Dynamic capabilities	–
Andreeva & Chaika, 2006 [31]	Functional capabilities	Core capabilities	Dynamic capabilities	–
Pavlou & Sawy, 2006 [32]	Underlying subprocess	First order capability	Second order capability	–
Wang & Ahmed, 2007 [33]	Capability	Core capability	Dynamic capability	–
Newey & Zahra, 2009 [34]	Core/operational capabilities	Dynamic capabilities	–	–
Ambrosini et al., 2009 [35]	*Regenerative* dynamic capabilities	Resource base	*Incremental* dynamic capabilities	*Renewing* dynamic abilities
Wójcik, 2015 [20]	Resources	Processes Routines	Lower-order capabilities (functional and operational)	Higher-order capabilities/strategic capabilities (core competencies)
Ceglinski, 2020 [36]	Dynamic capabilities	Core competencies	Core products/services	End products/services

In empirical research on strategic capabilities, authors increasingly choose different units of analysis. While research on strategic capabilities at the organizational level remains the most common, there is a growing interest in strategic capabilities at the individual level [21, 28]. In line with current trends towards the inclusion of context in capabilities research, strategic capabilities outside the organization are being considered, e.g., strategic capabilities at the industrial and country levels [7].

These assessment approaches have enriched the concept of strategic capabilities and have nuanced its understanding as an overarching multilevel paradigm. Despite a clear trend toward more detailed and specific approaches, there is still a desire to explore a general construct of strategic capabilities that incorporate several assessment approaches simultaneously and take into account the context of a particular capability.

A research team led by A. Cuervo-Cazurra, V. Newberry, and S. Park (Research Centre for Emerging Market Studies, Shanghai, China) developed an international approach to identify the main categories and subcategories of strategic capabilities of companies and to compare, in some terms, the results of studies for different countries and companies. The approach was validated in a large-scale study that compared different companies from 12 developing countries to identify which capabilities of leading multinational companies in emerging markets are strategically important for them [7]. The approach considered differences in strategic capabilities in the functional area and the focal unit of analysis for the capabilities at the administration, organization, industry, and country levels. Additionally, the study included assessments of strategic capabilities by depth (the degree of organizational knowledge improvement), breadth (the degree of diversity of organizational knowledge improvement), speed (the time required to achieve organizational knowledge improvement), and location (the geographic boundary where organizational knowledge improvement occurs) [7].

This approach identified five main categories responsible for the functional area: obtaining resources, product/service capabilities, operations and management, marketing, and managing the external environment. These categories include a broad list of subcategories of strategic capabilities, which can be expanded according to the specific capabilities of companies. Within this approach, companies are categorized using two criteria to capture their context: the industry of operation (low-tech, high-tech, or service companies) and the level of internationalization (domestic, export, or multinational companies). Therefore, the strategic capabilities of companies are investigated in a comprehensive and uniform way.

2. THE RESEARCH METHODOLOGY

In the empirical part of this study, the attention was focused on strategic capabilities as a driver of competitiveness of Russian domestic companies, Russian exporters, and multinational companies operating in the Russian market.

The main research question was formulated as follows: *What are the strategic capabilities underlying the competitiveness of the companies operating in the Russian market?* It was decomposed into three first-level questions:

1. What are the dominant strategic capabilities of Russian domestic companies?

2. What strategic capabilities dominate in the companies expanding their activities outside the local market (on an example of Russian exporters)?

3. What are the dominant strategic capabilities of multinational companies?

The methodology proposed by the international research group [7] was applied to study strategic capabilities. In addition to the five categories of strategic capabilities (obtaining resources, product/service capabilities, operations and management, marketing, and managing the external environment), we considered a new specific category (strategy). This category includes subcategories that could not be correlated with the existing categories of strategic capabilities (flexibility, long-term perspective and sustainability, balance price & quality, merger and acquisition, business portfolio management, etc.). Moreover, there is no unified approach to interpreting the relationship between strategy and strategic capabilities and no generally accepted opinion on which of them is primary [37]. Hence, the causal nature of the concepts does not prevent the emergence of a new category. The list of subcategories of strategic capabilities was expanded in accordance with the analysis data.

This study is based on semi-structured interviews with owners, top managers, and middle management representatives who correctly assess the components of the company's strategy and its competitiveness.

A total of 30 interviews were conducted with representatives of high-tech and low-tech industrial companies and service companies operating in the Russian market, which were distributed into 9 clusters by two criteria [7]: the industry of operation (high-tech, low-tech, or service companies) and the level of internationalization (domestic, export, or multinational companies) (Table 3).

Table 3

Characteristics of the selected companies

		Company category		
		Russian domestic companies	Russian exporters	Multinational companies
Industry type	*High-tech*	SUENKO JSC Laser Systems and Technologies LLC	VARTON JSC SPC ASPECT Sberbank PJSC	Merck BMW Group Toyota Motor Corporation
	Low-tech	SUEK JSC AISFER JSC Torgservis LLC (production, sales, and purchase of oil equipment)	AIC PROMAGRO LLC Kuzbassrazrezugol' OJSC	Japan Tobacco International GRAND VISION Kraft Heinz ECCO
	Services (*service*)	NPP Radintekh LLC Clean Town LLC Tsargrad LLC MC Medved LLC Prime Stomatologia LLC	Office Solutions LLC Novikov Group MILDBERRY R:TA Uchi.ru LLC	Ketchum Maslov Lotte Mazars

The empirical part of the study includes 10 Russian domestic companies, 10 Russian exporters, and 10 multinational companies. The manufacturing sector is represented by 16 companies and the service sector by 14 companies. Companies that use sophisticated technologies in their production or manufacture an innovative technically complex product were classified as high-tech companies. One company from the banking services sector, Sberbank PJSC, was classified as high-tech: it evolves as a modern ecosystem embracing modern technologies such as artificial intelligence (AI), robotic process automation (RPA), and Big Data, and penetrates other industries. Table A1 (see the Appendix) summarizes the main characteristics of the selected companies.

The interviews were subjected to content analysis to identify strategic capabilities in accordance with the categories and subcategories of strategic capabilities considered within the international approach [7]. Table A2 (see the Appendix) presents a fragment of the content analysis matrix with brief explanations. At the next stage, based on the content analysis matrix, the frequency of mentioning was calculated for each category and subcategory of strategic capabilities and for each type of companies (multinational, export, and domestic), and the results were ranked. The ranking procedure was carried out in several stages for each type of companies separately:

- the top 3 categories of strategic capabilities,
- the top 10 subcategories of strategic capabilities.

In the study, strategic capabilities were divided into 3 groups by the degree of their significance for companies:

- group 1, "very important" (70–100% of the companies, i.e., 7–10 companies out of 10);
- group 2, "important" (40–60% of the companies, i.e., 4–6 companies out of 10);
- group 3, "less important" (0–30% of the companies, i.e., 0–3 companies out of 10).

At the final stage of the research, the dominant categories and subcategories of strategic capabilities of Russian domestic companies, Russian exporters, and multinational companies were compared and analyzed.

3. THE DOMINANT STRATEGIC CAPABILITIES OF RUSSIAN COMPANIES

Among the categories of strategic capabilities mentioned by Russian domestic companies, the significance of *operations and management* stands out sharply; see Fig. 1. The second most important category is *product/service capabilities*, but the frequency of mentioning this category is comparable to *marketing* and *strategy* (both are ranked 3rd). Note a big difference in the significance of the top 3 categories of strategic capabilities of Russian companies.

The subcategories of strategic capabilities were analyzed to establish the focus of Russian companies when shaping their competitiveness.

The top 10 subcategories of strategic capabilities (Fig. 2) indicate that Russian domestic companies mostly prioritize human resources ("human assets... team" in MC Bear LLC). *Employee professionalism* belongs to the first group of strategic capabilities: it is very important in the activities of Russian domestic

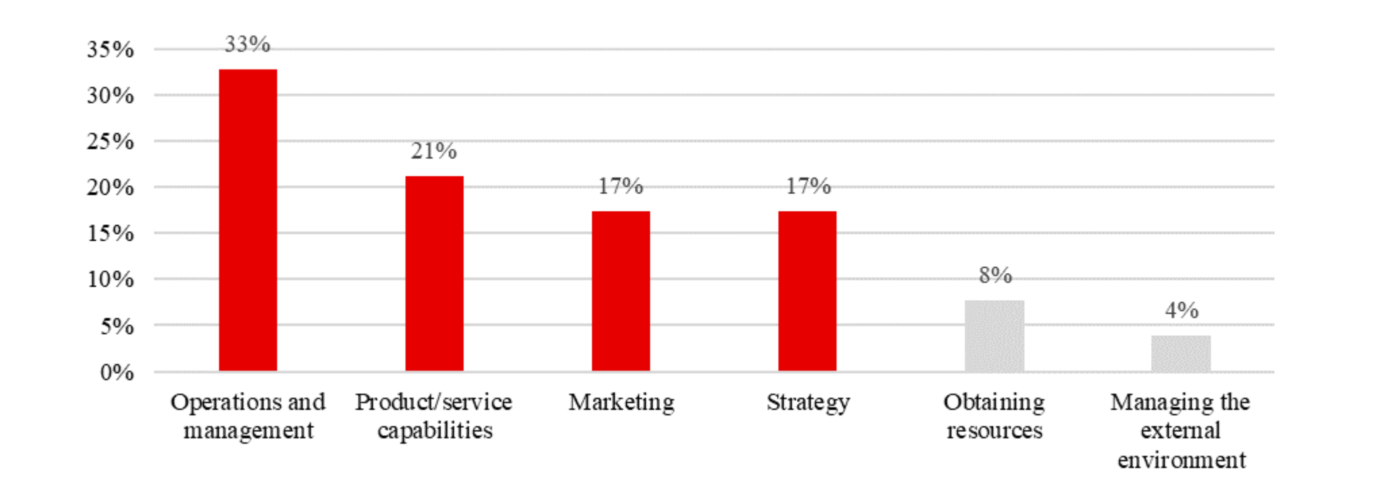

Fig. 1. The top 3 categories of strategic capabilities of Russian domestic companies.

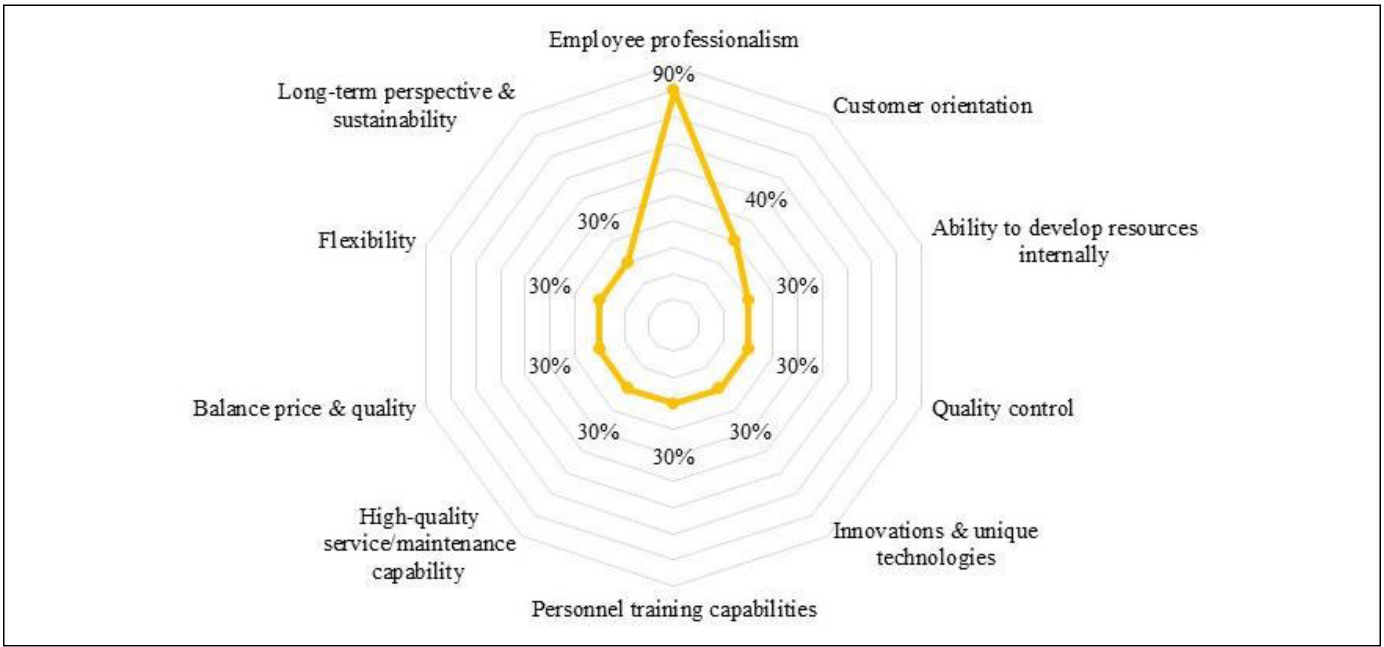

Fig. 2. The top 10 subcategories of strategic capabilities of Russian domestic companies.

companies. In separate cases, the interviewees emphasized the significance of *professional and flexible personal skills* depending on the specifics of the company's activity. Some of the capabilities classified as "less important" are also related to human resources. For example, 30% of the companies noted their developed *capabilities in personnel training*. They endeavor not to find but to nurture a professional: "we keep, train, and grow people all the way, beginning from scratch" (NPP Radintekh LLC); "we try to prepare our future employees starting from school; we follow them in their student years and strive to attract them in our company as much as possible" (SUENKO JSC).

Note that only *customer orientation* is at the lower boundary of the second group of strategic capabilities among Russian domestic companies. Thus, 40% of the Russian companies operating in the local market consider themselves customer-oriented. They apply an individual approach to customers: "...differentiated products, i.e., we have an individual approach to each customer" (NPP Radintekh LLC); "we try to optimize and adjust the program to the customer" (Clean City LLC). However, representing only one side of the concept of customer orientation, such an approach cannot predetermine the entire activity of the company as customer-oriented.

The third group includes such specific strategic capabilities as *the ability to develop resources internally* ("the full cycle of production...own raw material base" (AISFER JSC); "a large number of reserves, i.e., deposits" (SUEK JSC)) and *balance price & quality* ("a reasonable price/quality ratio" (Clean City LLC)). Their emergence was dictated by the historical development of the country and its resource dependence, which is gradually being decreased through realizing the existing potential in railway transport, agriculture, the military-industrial complex, and the digital economy. Another reason is the general state of the economy, where the cost of a product or service comes to the fore: on the one hand, companies strive for high margins and low costs; on the other, consumers seek a price corresponding to their standard of living.

According to the aforesaid, Russian domestic companies build their activities based on human capital and its potential, i.e., they mostly focus on a single strategic capability (*employee professionalism*).

4. THE DOMINANT STRATEGIC CAPABILITIES OF RUSSIAN EXPORTERS

The activities of Russian exporters are based on strategic capabilities related to the top 3 categories: *operations and management*, *marketing*, and *product/service capabilities* (Fig. 3). The companies focus on the above categories almost equally, mentioning a large number of diverse subcategories.

The next level (the analysis of subcategories of strategic capabilities) allowed us to identify the top 10 subcategories in the strategic capabilities of the Russian exporters (Fig. 4). For these companies, *operations and management* is ranked 1st; despite this fact, the most important subcategory of strategic capabili-

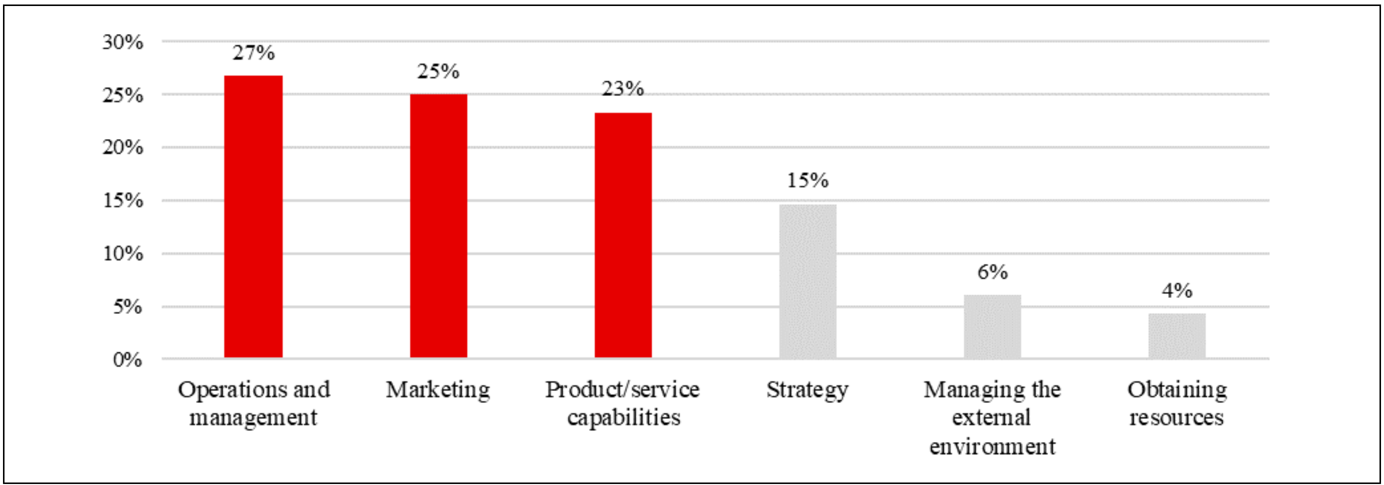

Fig. 3. The top 3 categories of strategic capabilities of Russian exporters.

Fig. 4. The top 10 subcategories of strategic capabilities of Russian exporters.

ties is *quality control* from *product/service capabilities*: its significance echoes the results for the Russian domestic companies.

The top 10 strategic capabilities also included *employee professionalism* and *strong top management*: their significance was noted by 70% and 60% of the interviewees, respectively. Russian exporters emphasize the importance of human capital in general and, moreover, focus on the level of management training.

The strategic capabilities of the Russian exporters belonging to the second group are the most diverse compared to the results obtained for the Russian domestic companies. Every second company (i.e., 50% of the Russian exporters) noted the significance of such strategic capabilities as *innovations & unique technologies*, *production management*, *corporate brand and reputation*, *customer orientation*, *market and trend understanding*, *relationship capabilities*, and *flexibility*.

Thus, the dominant strategic capabilities in the activities of the Russian exporters are as follows: *quality control*, *employee professionalism*, and *strong top management*. Nevertheless, generally speaking, the top 10 subcategories of strategic capabilities have high importance for Russian exporters. This approach is due to their desire to increase competitiveness in new geographical markets and to follow the rules of economic struggle in the international arena.

5. THE DOMINANT STRATEGIC CAPABILITIES OF MULTINATIONAL COMPANIES

According to the analysis results, multinational companies are attentive to both *management* and *marketing*, which are ranked 1st and 2nd, respectively, in the top 3 categories of their strategic capabilities (Fig. 5). *Product/service capabilities* are the third most important category for multinational companies.

The next deeper level of the analysis allowed us to identify the top 10 subcategories of the strategic capabilities of multinational companies (Fig. 6). The most important (the first group) strategic capabilities of multinational companies include *quality control* and *employee professionalism*. The high significance of quality for international competition was demonstrated back in the 1980s by the Japanese Economic Miracle: Japanese cars superior to American counterparts in terms of reliability and cost were introduced to the US market and won the struggle for the market at that time. Quality also correlates with resource scarcity, a global problem generating a priority task for the world economy (resource saving).

The quality management system and innovative potential of a company are directly related to human capital. People are becoming the key asset of companies. Multinational companies actively invest in human resources, hiring competent staff (*employee professionalism*). Indeed, it is impossible to manufacture a high-quality product without well-prepared, trained personnel: "the company is focused....on qualified specialists" (Merck Group); "... on the people, i.e., on the expertise of our workers and on our methodology, which actually depends on the qualification of people" (Mazars).

The second group—"important strategic capabilities"—includes capabilities from *marketing*: *understanding local/foreign customer needs*, *corporate brand and reputation*, and *understanding market and trends*. (The latter two subcategories were mentioned almost as many times as *marketing*.) Today, the orientation of the global market to customers and satisfaction of their needs and desires are clear: customer-oriented companies are 60% more profitable than

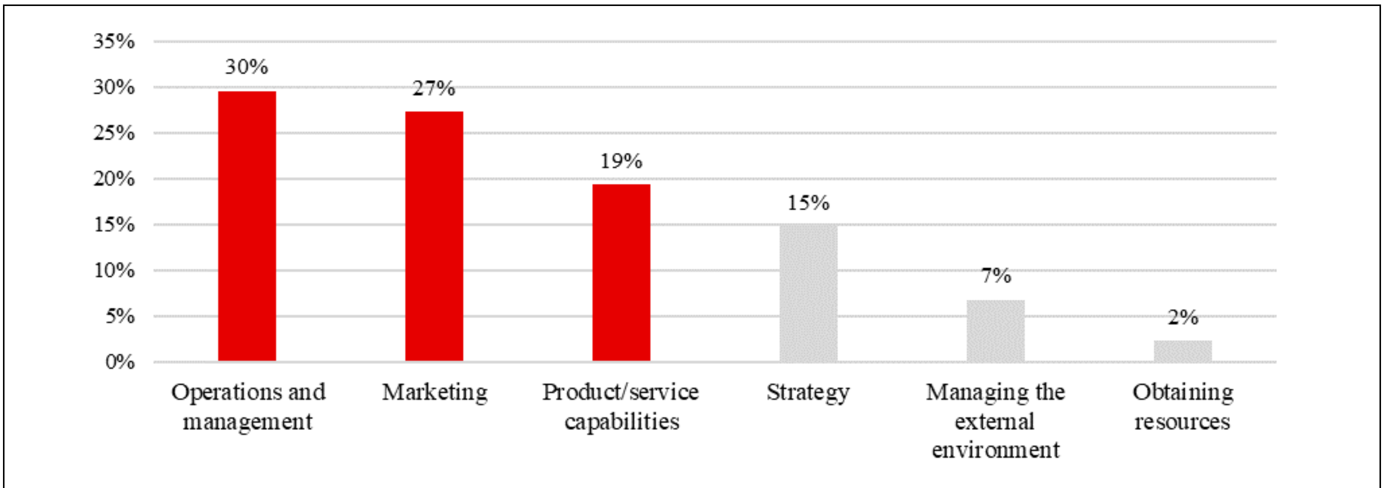

Fig. 5. The top 3 categories of strategic capabilities of multinational companies.

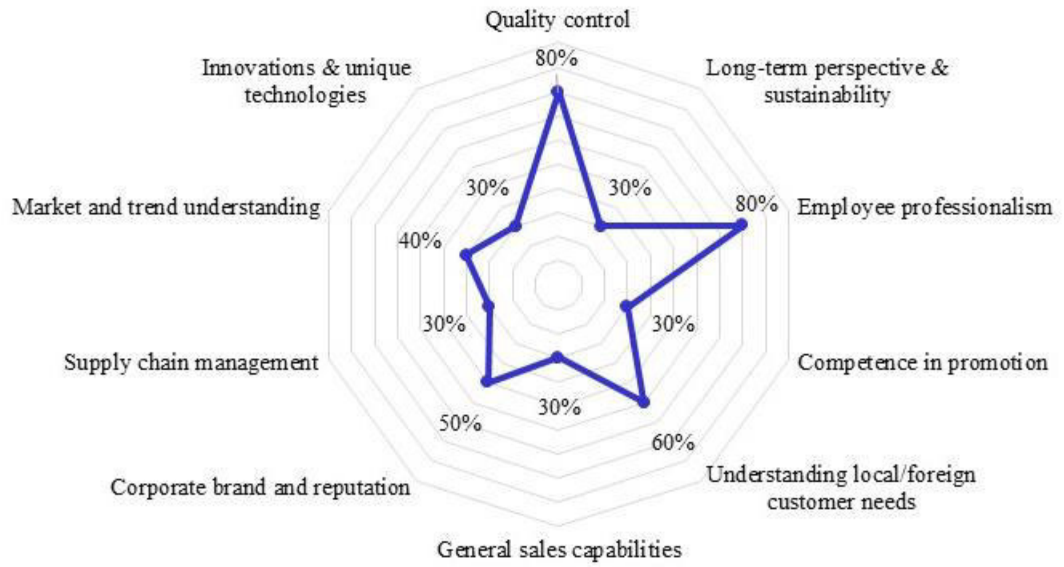

Fig. 6. The top 10 subcategories of strategic capabilities of multinational companies.

those without customer emphasis in their work; see [Forbes, 2019][1]. In addition, many large multinational companies associate their brand with quality and are therefore responsible for their reputation in the market. Finally, constant market research, consideration of current trends, and understanding consumers serve to adjust the level of quality and necessary costs.

The third group—less important strategic capabilities—contains the remaining ones from the top 10 subcategories: *innovations & unique technologies*, *supply chain management*, *general sales capabilities*, *competence in promotion*, and *long-term perspectives and sustainability*.

As a result, we identified five dominant strategic capabilities in the activities of multinational companies:

– *quality control*,
– *employee professionalism*,
– *understanding local/foreign customer needs*,
– *corporate brand and reputation*,
– *understanding market and trends.*

Among the strategic capabilities of multinational companies, the second half of the top 10 subcategories have moderate significance compared to the first half. Nevertheless, thanks to them, the company can remain agile and precise in the economic struggle. The competitive approach of multinational companies can be considered a benchmark for those seeking to enter the international arena and increase the level of competitiveness in the markets.

6. RUSSIAN DOMESTIC COMPANIES, RUSSIAN EXPORTERS, AND MULTINATIONAL COMPANIES: A COMPARISON OF STRATEGIC CAPABILITIES

The top 3 categories of the strategic capabilities of Russian domestic companies, Russian exporters, and multinational companies were compared. According to the comparison results, for all groups of companies, *operations and management* is ranked 1st. For Russian exporters and multinational companies, *marketing* is the second most important source of strategic capabilities; for Russian domestic companies, however, *product/service capabilities* are more important than *marketing*, which is (in turn) commensurate with *strategy*. Note a significant gap in the frequency of mentioning the categories of strategic capabilities by different groups of companies. This means a greater variety of subcategories of strategic capabilities that are the most important for Russian exporters. Thus, companies operating in the Russian market should look for strategic capabilities to improve their competitiveness in three directions corresponding to the dominant categories of strategic capabilities of the companies under consideration:

- *operations and management*,
- *marketing*,
- *product/service capabilities.*

[1] Morgan, B., 50 Stats That Prove the Value of Customer Experience, *Forbes*, 2019. URL: https://www.forbes.com/sites/blakemorgan/2019/09/24/50-stats-that-prove-the-value-of-customer-experience/?sh=53d0e67b4ef2 (Accessed August 24, 2022.)

Based on the comparison results of the top 10 subcategories of strategic capabilities (Fig. 7), we found common points in the competitiveness of Russian domestic companies, Russian exporters, and multinational companies as well as differences characterizing their activities depending on the level of internationalization.

For example, multinational companies form their competitiveness by mainly focusing on strategic capabilities related to quality, people, and marketing. These strategic capabilities correlate with the dominant capabilities of Russian domestic companies and Russian exporters through a common feature: concentration on people and their skills. *Employee professionalism* can be considered a dominant strategic capability of companies competing in the Russian market.

Quality control is a dominant strategic capability only for multinational companies and Russian exporters, being less important for Russian domestic companies. In Russia, state quality control disappeared with the transition to the market economy; as a result, in the early 1990s, the market was flooded with counterfeit products from all over the world. After 20 years, mandatory certification of products was introduced to ensure compliance with safety requirements but not quality. Quality control was under the control of business, or rather money or demand. The National Certification System began to operate in Russia in 2017, establishing the requirements of voluntary quality standards. In some situations, it is impossible to manage without them, which makes such standards mandatory. The year 2022 showed the significant success of system import substitution in the country. Companies in Russia turned out to be ready to supply analogs of foreign goods of different quality. Therefore, quite surprisingly, Russian companies do not emphasize the significance of quality control. Voluntary and mandatory quality standards are still applied in the country, and consequently, this area is not perceived by managers as the one that they should stress in business on their own initiative. Also, probably, managers of domestic companies do not realize the importance of this capability in building the competitiveness of the company or its relevance for carrying out activities in the core industry. Against this background, some Russian exporters raise their strategic capabilities to the level of

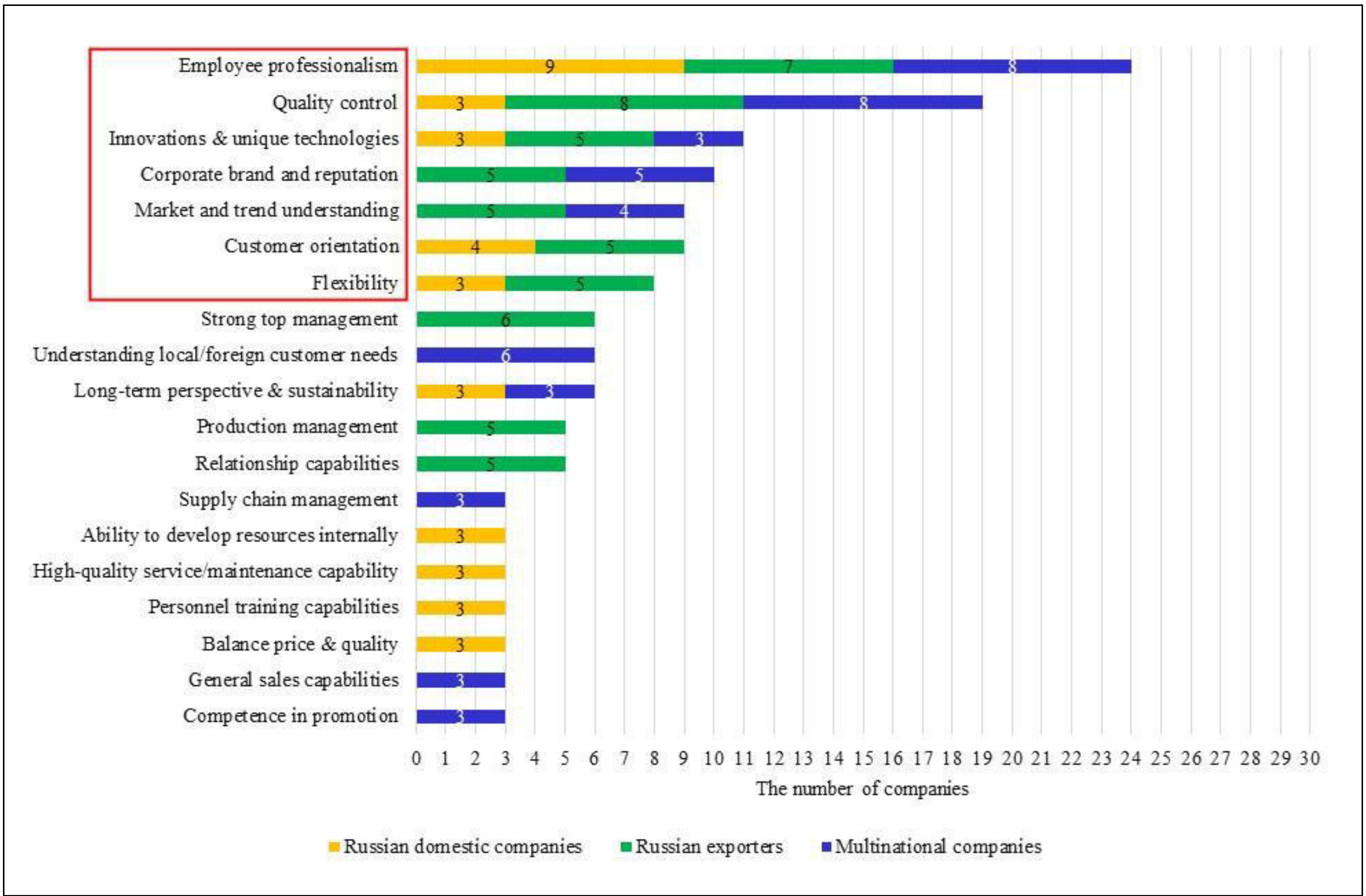

Fig. 7. The strategic capabilities of Russian domestic companies, Russian exporters, and multinational companies: a comparison of the top 10 subcategories.

competitiveness in global markets and develop this subcategory of strategic capabilities, making it a dominant one in the Russian market.

In contrast to other groups of companies, Russian exporters have a high significance of *strong top management*, which generally characterizes Russian business in the context of prevailing rigid vertical organizational structures. Another reason is the need of companies for a strong leader, especially in a very turbulent external environment and their desire for international expansion.

The significance of marketing brought to a high level is common only for multinational companies. Russian domestic companies still partially realize the importance of market research, understanding the consumer, and creating the face of the company. Meanwhile, Russian exporters are actively working to develop and increase the recognition of their brand and are trying to give a good account of their product/service in new markets. This fact brings Russian exporters closer to the international style of doing business, but the corresponding strategic capabilities cannot be considered dominant so far.

The number of dominant subcategories in strategic capabilities differs between the groups of companies, which directly affects their competitive style. In Russian domestic companies, only one strategic capability has a high significance: they apply pinpoint strikes in competition. In Russian exporters, three dominant strategic capabilities were identified, and the other capabilities have a small gap in importance with them. Russian exporters are gaining power to be able to compete successfully in the local (Russian) market with players that have been operating internationally for a long time. Finally, multinational companies can be treated as a benchmark, as they already have experience in various markets and successfully operate in the Russian market. Their set of strategic capabilities can be considered balanced; together with rich experience, this fact makes multinational companies dangerous opponents in the competitive struggle. Thanks to their five dominant strategic capabilities, multinational companies can successfully overcome emerging obstacles and damage competitors.

As a result, multinational companies demonstrate a set of strategic capabilities necessary to be competitive in any geographical market. Russian domestic companies actively develop only human capital out of this set. Russian exporters overestimate the strategic nature of their capabilities. They are an intermediate link in the process of transforming a company from domestic to multinational. By transforming their activities, Russian exporters retain the features of local players but evolve towards multinational companies by focusing on strategic capabilities that are important for both local and global competition. The dominant subcategories of the strategic capabilities of Russian domestic companies, Russian exporters, and multinational companies that form their competitiveness in the Russian market are as follows:

- *employee professionalism*,
- *quality control*,
- *innovations & unique technologies*,
- *corporate brand and reputation*,
- *understanding local/foreign customer needs*,
- *customer orientation*,
- *flexibility*.

Thus, the strategic capabilities of Russian domestic companies, Russian exporters, and multinational companies were compared to identify the set of strategic capabilities forming their competitiveness. According to the results, the level of internationalization of the company affects the set of its dominant strategic capabilities, which changes under the conditions of the external environment in the market.

This study contributes makes to the strategic capability literature, both theoretically and practically. On the one hand, it expands the field of strategic capabilities research by providing empirical results on strategic capabilities in a specific location and accumulating approaches to their study. On the other hand, the main findings of the study have applied significance: they can help top management maintain and improve the competitiveness of their companies when operating in the Russian market and when entering new foreign markets. Also, the findings indicate which strategic capabilities managers should invest in for higher competitiveness of their companies and better performance in the competitive game in the Russian market.

CONCLUSIONS

The theoretical part of this study has considered the main conceptual provisions of strategic capabilities. Accumulating unique resources, core competencies, knowledge, and abilities that a company should use most effectively in response to the challenges of the external and internal environment, strategic capabilities underlie long-term competitive advantages of the company. Researchers apply various approaches to assessing strategic capabilities based on their differences in the functional area, the capability hierarchy, and the focal unit of analysis.

The empirical part of this study has considered 30 high-tech, low-tech, and service companies operating in the Russian market. Their strategic capabilities have been investigated to determine the dominant categories and subcategories of strategic capabilities of the companies when forming their competitiveness.

According to the subcategory analysis, the competitiveness of companies in the Russian market is based on the following set of strategic capabilities: *employee professionalism*, *quality control*, *innovations & unique technologies*, *corporate brand and reputation*, *understanding local/foreign customer needs*, *customer orientation*, and *flexibility*. Three areas have been identified to find other strategic capabilities that can improve the competitiveness of companies: *operations and management*, *marketing*, and *product/service capabilities.*

The results of this study are of academic value for further research on the strategic capabilities of companies: they expand the subject domain with new empirical data. The practical significance of the results for the business community consists in the most significant sources to increase the competitiveness of companies operating in the Russian market as well as in new foreign markets.

APPENDIX

Table A1

Information about the companies selected for the empirical study

Company name	Sector of economic activity	Year of foundation	Country of origin	Revenue 2019, USD[a]	Number of markets
Russian domestic companies					
SUENKO JSC	Electricity	2002	Russia	264.336 million	1
Laser Systems and Technologies LLC	Printing presses, other offset printing machines, and other equipment	2008	Russia	n/a	1
SUEK JSC	Coal industry	2001	Russia	7.547 million	1
AISFER JSC	Food industry	2011	Russia	21.267 million	1
Torgservis LLC	Production, trade in oil equipment	n/a	Russia	n/a	1
NPP Radintekh LLC	Research and development services	2016	Russia	525 116	1
Clean Town LLC	Advertising industry	2005	Russia	116 938	1
Tsargrad LLC	Security services	1996	Russia	1.178 million	1
MC Medved LLC	Healthcare	2014	Russia	1.284 million	1
Prime Stomatologia LLC	Healthcare	2005	Russia	2.151 million	1
Russian exporters					
Varton	LED production	2009	Russia	12.647 million	4
JSC SPC ASPECT	Manufacture of instruments and devices for measurement, testing, and navigation	1991	Russia	24.087 million	32
Sberbank PJSC	Banking and financial services	1991	Russia	13.972 million	22
AIC Promagro LLC	Pork production, food industry	2014	Russia	101.169 million	6
Kuzbassrazrezugol’ OJSC	Coal industry	2003	Russia	2.433 billion	n/a
Office Solutions LLC	Wholesale trade in office furniture	1995	Russia	59.5 million	5
Novikov Group	Catering	1991	Russia	n/a	8
MILDBERRY	Design industry	1993	Russia	1.641 million	n/a
R: TA	Advertising industry	n/a	Russia	n/a	n/a
Uchi.ru LLC	Software industry	2012	Russia	16.058 million	n/a
Multinational companies					
Merck Group (Russian subsidiary is Sigma-Aldrich)	Chemical and pharmaceutical industry	1668	Germany	18.145 million	66

Table A1 (continued)

BMW Group	Automotive industry	1916	Germany	116.833 billion	8
Toyota Motor Corporation	Automotive industry	1937	Japan	278.460 billion	190
Japan Tobacco International	Tobacco industry	1949	Japan	20.268 billion	130
GRAND VISION	Optical industry	1891	Netherlands	4.535 million	40
Kraft Heinz	Food industry	1869 (M&A Kraft Foods and H.J. Heinz in 2015)	USA	24.98 billion	50
ECCO	Footwear industry	1963	Denmark	1.527 million	99
Ketchum (Russian subsidiary is Ketchum Maslov)	Consulting, Marketing Communications	1923	USA	n/a	8
Lotte	Conglomerate	1967	South Korea	15.252 billion	8
Mazars	Audit, consulting	1945	France	2 billion	91
[a] Exchange rates are as of December 31, 2019: USD 1 = EUR 0.890155 USD 1 = JPY 109.593831 USD 1 = KRW 1 153.889977 USD 1 = RUB 62.271942					

Table A2

One fragment of the content analysis matrix

	Industry type	Low-tech				…	TOTAL		
		Company name				…			
		Kraft Heinz		AIC Promagro LLC		…			
No.	Category / Subcategory of strategic capabilities	Code	Description (quote)	Code	Description (quote)	…	Qty of "1+"	Qty of "11+"	Qty of "111+"
4	Marketing						Sum of all cells coded "1+" within category 4	...	...
4.1	General sales capabilities	0		0		…	...	...	...
4.2	Market and trend understanding	111+	Knowledge, consumer expertise (p. 92)	11+	Examine the requirements of a network client (p. 112)	…	Sum of all cells coded "1+" in row 4.2	...	...
4.3	Understanding local/foreign customer needs	0		0		…	...	...	...
4.4	Corporate brand and reputation	111+	Strong brand (p. 73)	11+	Development and promotion of own brand (p. 26)	…	...	...	...
4.5	Customer orientations	0		11+	Adapt to individual customer characteristics (pp. 136 and 137)	…	...	...	...
...	...	...	...	...	...	…	...	...	...

Code Legend:
"1+"—Russian companies operating in the domestic market;
"11+"—Russian companies operating in both domestic and foreign markets;
"111+"—multinational companies represented in the Russian market.

REFERENCES

1. Kumar, V., Simon, A., and Kimberley, N., Strategic Capabilities Which Lead to Management Consulting Success in Australia, *Management Decision*, 2000, vol. 38, no. 1/2, pp. 24–35.
2. Barney, J.B., Firm Resources and Sustained Competitive Advantage, *Journal of Management*, 1991, vol. 17, no. 1, pp. 99–120.
3. Song, M., Strategic Capabilities of Emerging Market Firms, FIU Electronic Theses and Dissertations, 2021.
4. Drucker, P., *Management Challenges for the 21st Century*, Butterworth–Heinemann, 1999.
5. Lenz, R.T., Strategic Capability: A Concept and Framework for Analysis, *Academy of Management: The Academy of Management Review*, 1980, vol. 5, no. 2, pp. 225–324.
6. Johnson, G., Scholes, K., and Whittington, R., *Exploring Corporate Strategy*, Pearson Education Limited, 2008.
7. Cuervo-Cazurra, A., Newburry, W., and Park, S., *Building Strategic Capabilities in Emerging Markets*, Cambridge University Press, 2020.
8. Robert, M., *Strategy Pure and Simple: How Winning CEOs Outthink Their Competition*, McGraw-Hill: Sydney, 1993.
9. Teece, D.J., Pisano, G., and Shuen, A., Dynamic Capabilities and Strategic Management, *Strategic Management Journal*, 1997, vol. 18, no. 7, pp. 509–533.
10. Prahalad, C.K., Developing Strategic Capability: An Agenda for Top Management, *Human Resource Management*, 1983, vol. 22, no. 3, pp. 237–254.
11. Hayes, R.H., Pisano, G.P., and Upton, D.M., *Strategic Operations: Competing Through Capabilities*, Sydney: The Free Press, 1996.
12. Cool, K., Costa, L.A., and Dierickx, I., Constructing Competitive Advantage, in *Handbook of Strategy and Management*, Pettigrew, A., Thomas, H., and Whittington, R., Eds., Thousand Oaks, CA: Sage, 2002.
13. Teece, D.J., The Nature, Sources and Consequences of Firm Differences in the Early History of the Semiconductor Industry, *Strategic Management Journal*, 2000, vol. 21, no. 10-11, pp. 1017–1042.
14. King, A.W. and Zeithaml, C.P., Competencies and Firm Performance: Examining the Causal Ambiguity Paradox, *Strategic Management Journal*, 2001, vol. 22, no. 1, pp. 75–99.
15. Kay, J., The Structure of Strategy, *Business Strategy Review*, 1993, vol. 4, no. 2, pp. 17–37.
16. Johne, A., *Market Driven Strategy: Processes for Creating Value*, New York: The Free Press, 1992.
17. Stalk, G., Evans, P., and Shulman, L.E., Competing on Capabilities: The New Rules of Corporate Strategy, *Harvard Business Review*, 1992, vol. 70, no. 2, pp. 57–69.
18. Day, G., The Capabilities of Market-Driven Organizations, *Journal of Marketing*, 1994, vol. 58, no. 4, pp. 37–52.
19. Helfat, C.E. and Peteraf, M.A., The Dynamic Resource-Based View: Capability Lifecycles, *Strategic Management Journal*, 2003, vol. 24, no. 10, pp. 997–1010.
20. Wójcik, P., Exploring Links Between Dynamic Capabilities Perspective and Resource-Based View: A Literature Overview, *International Journal of Management and Economics*, 2015, no. 45, pp. 83–107.
21. Schilke, O., Hu, S., and Helfat, C., Quo Vadis, Dynamic Capabilities? A Content-Analytic Review of the Current State of Knowledge and Recommendations for Future Research, *Academy of Management Annals*, 2018, vol. 12, no. 1, pp. 390–439.
22. Hafeez, K., Zhang, Y.B., and Malak, N., Determining Key Capabilities of a Firm Using Analytic Hierarchy Process, *International Journal of Production Economics*, 2002, vol. 76, no. 1, pp. 39–51.
23. Inan, G.G. and Bititci, U.S., Understanding Organizational Capabilities and Dynamic Capabilities in the Context of Micro Enterprises: A Research Agenda, *Procedia - Social and Behavioral Sciences*, 2015, no. 210, pp. 310–319.
24. Van Looy, A., De Backer, M., and Poels, G.A., Conceptual Framework and Classification of Capability Areas for Business Process Maturity, *Enterprise Information Systems*, 2012, pp.1–37.
25. LeanIX. Business Capability Map. URL: https://www.leanix.net/en/resources/business-capability-map-library/#/Default. (Accessed August 24, 2022.)
26. Pandza, K., Horsburgh, S., Gorton, K., and Polajnar, A., A Real Options Approach to Managing Resources and Capabilities, *International Journal of Operations & Production Management*, 2003, vol. 23, no. 9, pp. 1010–1032.
27. Vesalainen, J. and Hakala, H., Strategic Capability Architecture: The Role of Network Capability, *Industrial Marketing Management*, 2014, vol. 43, no. 6, pp. 938–950.
28. Collis, D., How Valuable are Organizational Capabilities?, *Strategic Management Journal*, 1994, vol. 15, no. S1, pp. 143–152.
29. Danneels, E., The Dynamics of Product Innovation and Firm Competences, *Strategic Management Journal*, 2002, no. 23, pp. 1095–1121.
30. Winter, S.G., Understanding Dynamic Capabilities, *Strategic Management Journal*, 2003, no. 24, pp. 991–995.
31. Andreeva, T. and Chaika, V., Dynamic Capabilities: What They Need to Be Dynamic?, *Discussion paper of the Institute of Management*, St. Petersburg State University, 2006.
32. Pavlou, P.O.A. and Sawy, E.L., Decomposing and Leveraging Dynamic Capabilities, *Working Paper of the Anderson Graduate School of Management*, University of California, Riverside, 2006.
33. Wang, C.L. and Ahmed, P.K., Dynamic Capabilities: A Review and Research Agenda, *International Journal of Management Reviews*, 2007, vol. 9, no. 1, pp. 31–51.
34. Newey, L.R. and Zahra, S.A., The Evolving Firm: How Dynamic and Operating Capabilities Interact to Enable Entrepreneurship, *British Journal of Management*, 2009, no. 20, pp. 81–100.
35. Ambrosini, V. and Bowman, C., What Are Dynamic Capabilities and Are They a Useful Construct in Strategic Management?, *International Journal of Management Reviews*, 2009, vol. 11, no. 1, pp. 29–49.
36. Ceglinski, P., The Relations between Dynamic Capabilities and Core Competencies on the Case of Polish Companies, *Administrative Sciences*, 2020, vol. 10, no. 48, pp. 1–10.
37. Johannesson, J. and Palona, I., The Dynamics of Strategic Capability, *International Business Research*, 2010, vol. 3, no. 1, 3–12.

This paper was recommended for publication by V.V. Klochkov, a member of the Editorial Board.

Received December 30, 2022, and revised June 27, 2023. Accepted August 30, 2023.

Author information

Guseva, Natal'ya Igorevna. Dr. Sci. (Soc.), PhD in Management, National Research University Higher School of Economics, Moscow, Russia
✉ natguseva@hse.ru
ORCID iD: https://orcid.org/0000-0002-2907-0737

Trubnikova, Olesya Yur'evna. Postgraduate, National Research University Higher School of Economics, Moscow, Russia
✉ otrubnikova@hse.ru
ORCID iD: https://orcid.org/0000-0003-2298-9107

Cite this paper
Guseva, N.I., Trubnikova, O.Yu., Strategic Capabilities as a Driver of Competitiveness: a Comparison of Russian and Global Companies. *Control Sciences* **5,** 41–55 (2023). http://doi.org/10.25728/cs.2023.5.4

Translated into English by *Alexander Yu. Mazurov*, Cand. Sci. (Phys.–Math.), Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
✉ alexander.mazurov08@gmail.com

DOI: http://doi.org/10.25728/cs.2023.5.5

MODELING SOCIAL ATTITUDE TO INTRODUCING EPIDEMIC SAFETY MEASURES IN A PANDEMIC[1]

I.M. Azhmukhamedov[2] and D.A. Machueva[3]

[2]Astrakhan State University, Astrakhan, Russia
[3]Grozny State Oil Technical University, Grozny, Russia

✉ aim_agtu@mail.ru, ✉ ladyd_7@mail.ru

Abstract. The COVID-19 pandemic is a global human-scale emergency that has caused many negative effects. To mitigate them, it is necessary to take competent and well-founded organizational measures. Considering infectious diseases from a mathematical point of view allows solving problems in various spheres of society, studying possible scenarios, identifying epidemiological evolution patterns, and proposing intervention strategies and epidemic control actions. This paper presents a mathematical model for forecasting opinion dynamics on various socially significant issues, in particular, on the introduction of epidemic safety measures in a pandemic. The model reflects the process of information exchange considering the content of disseminated information and the communicative properties of the social system and its elements (connectivity, susceptibility, and sociability).

Keywords: social system, process modeling, informational control, distribution of opinions, anti-epidemic measures.

INTRODUCTION

In the last few years, there has been a concept in the information space whose social significance seems difficult to exaggerate: pandemic. It is defined as the intensity of an epidemic process characterized by the mass spread of an infectious disease in several countries or even continents.

The COVID-19 infection faced by mankind has several features causing numerous difficulties in diagnosis and treatment, as well as in forecasting the development and making timely organizational and managerial decisions, particularly at the regional level.

The key problem is the absence of accumulated statistical data. The frequent mutation of this virus and asymptomatic and mild cases considerably complicate analytical studies [1].

In addition to the epidemiological characteristics of a pandemic and the socio-demographic structure of the population, the indicators of pandemic development depend on many other factors, especially, on the prevention and treatment measures taken by the healthcare system, the attitude to them in society, and the social position of the majority.

Despite these difficulties, attempts are undertaken to develop a mathematical framework and software to model the real epidemiological and socio-economic situation in a region and to forecast its dynamics and the consequences of certain planned measures. Considering infectious disease models allows conducting computational experiments, studying possible scenarios, identifying epidemiological evolution patterns, and proposing intervention strategies and epidemic control actions.

1. MODELS TO DESCRIBE AND FORECAST EPIDEMICS

Modeling may aim at short- or long-term forecasting of the epidemiological situation, assessing the nature and dynamics of the infection spread, identifying key time periods (the peaks of incidence, reaching a

[1] This study was supported by the fundamental research project "Methodological Fundamentals of Assessment and Control of Regional Complex Security" within the Development Program of Astrakhan State University for 2021–2030.

plateau, and decay), and testing of anti-epidemic measures. Depending on the problems formulated, researchers prefer one or another type of models [2, 3].

There are analysis methods based on considering epidemic processes in continuous and discrete time, at the level of interaction between population groups separated by different attributes and individuals [4].

A “gold standard” in describing epidemics is the SIR chamber model and its modifications. According to this concept, a population is divided into groups (chambers) depending on the state (the stage of the disease). Different modifications of this model include three to seven groups are follows: *S* (Susceptible), *E* (Exposed, in the incubation period), *I* (Infected, with symptoms), *R* (Recovered), *H* (Hospitalized, with severe disease dynamics), *C* (Critical, requiring a lung ventilator), and *D* (Deaths).

Over time, the status of individuals changes and they pass from one group to another. Thus, nowadays, there exists a whole family of models of this class with different sets of chambers: SIR, SIS (without immunity production), SIRS (with temporary immunity), SEIR, SIRD, SEIRD, and SEIR-HCD [5].

The mathematical model represents a system of nonlinear partial differential equations with initial and boundary-value conditions [6]. The parameters of SIR models determine the frequency and probability of transitions between groups, i.e., the rates of recovery and re-infection, the frequency of symptoms, the probability of hospitalization, the disconnection of the lung ventilator, etc.

These indicators are calculated based on the demographic and geographic characteristics of the territory under consideration (country, region, or city), ideally considering multiple factors (the type of disease, the probability of virus mutations, population density, mobility, age, average immunity level, and climate) [7].

The main problem of chamber models is that their output results have high sensitivity to variations in their input parameters. The lack of reliable statistical data significantly reduces the quality of forecasting.

Agent-based models are more detailed than chamber models. In such models, each participant is considered individually taking into account its social links. Agents also go through disease stages, but all transitions are modeled at the individual level rather than at the group level. As a result, population heterogeneity can be studied in terms of different personal characteristics, e.g., basic health level and the number of social contacts. To succeed in developing a model of this class, it is necessary to represent the structure of the modeled system and to simulate the actions of agents with sufficient accuracy [8].

In the class of so-called reaction-diffusion epidemiological models, the spread of viral infection is described by a system of heat conduction equations. They assume the non-uniform spatial distribution of participants but an instantaneous virus transmission without any incubation period [9].

Regression and time series analysis methods are well known to forecast morbidity (incidence rate). In particular, we note ARIMA, an integrated autoregression model. It is a class of parametric models describing nonstationary time series. However, the absence of reliable statistics for the past periods significantly restricts the application of this approach.

2. INFORMATIONAL INFLUENCE AND CONTROL IN EPIDEMIOLOGICAL EVOLUTION

To improve the efficiency and adequacy of models, it is necessary to consider the influence of administrative measures intended to limit the virus spread.

Such measures include social distancing, self-isolation, cancellation of mass events, blocking of transportation flows, remote production and training, and wearing masks and gloves in public places. They reduce the quality of life, worsen the psychological state of people, and may cause negative reactions [10–12].

The efficiency of quarantine measures largely depends on the discipline of the population [13]. Therefore, they should be introduced considering special mathematical models-based assessments with additional input data related to social processes rather than to medical aspects.

Several years before the global spread of COVID-19, models were developed to determine possible human behavior under various countermeasures for diseases [14, 15]. The corresponding research aimed to develop disease control strategies and assess different intervention actions [16]. The studies of behavioral incentives identified an obvious correlation between epidemiological and economic indicators [17]. Thus, when determining the influence of various restrictive measures on epidemiological evolution, it is necessary to consider the degree of their public support or rejection.

Domestic and foreign researchers investigate the processes of information dissemination in social systems and the influence of various factors on the formation of opinions (beliefs and attitudes) in society.

The general patterns of opinion dynamics (agreement and convergence, divergence and polarization) were considered and justified in [18, 19]. The factors determining the value and attractiveness of information, various personal characteristics reflecting the internal state and responsible for the external behavior of people, as well as the rules of information transfer in communication processes were identified and assessed in [20, 21].

Many publications have been devoted to the problems of informational control, i.e., planning and organization of indirect and implicit influences to distribute information that inclines social system participants to choose a required line of behavior [22]. Different strategies have been proposed to regulate public opinion on significant issues (in particular, the identification and use of critical points of social networks, the so-called "opinion leaders" [23, 24]).

The models of informational influence proceed from various basic theoretical assumptions and hypotheses: linear threshold models, independent cascade models, models with biological and thermodynamic analogies ("contamination" and "magnetization"), cellular automata models, Markov chain models, and game-theoretic models. Each model mentioned has specific advantages in a particular sphere; at the same time, mathematical models are often based on several assumptions, which can complicate their use in the applied research of real social interactions [25–27].

In what follows, we propose a mathematical model of information exchange in a social system based on aggregated indicators (probability distributions). The model considers the content of the information block disseminated and some communication properties of the social system and its elements (the number of connections, susceptibility, and sociability).

This model can be used to forecast the distribution of opinions in society on the disseminated information about the introduction of epidemic safety measures in a pandemic as well as to elaborate managerial decisions toward increasing the public awareness of the importance of anti-epidemic measures.

3. MATHEMATICAL MODEL TO FORECAST OPINION DYNAMICS IN SOCIETY

The population living in a certain region is a social system characterized by connectivity (the average number of social contacts), sociability, and susceptibility of its elements (individuals) to some information. Sociability is the desire to share the information received, while susceptibility is the inclination of an individual to change his (or her) point of view under the influence of others.

The model parameters related to the individual's characteristics cannot be quantified precisely. Therefore, these data are formalized by introducing a distribution with a value set consisting of three or five categories as follows:

$$\{\text{low } (L);\ \text{moderate } (M);\ \text{high } (H)\}\,; \qquad (1)$$

$$\{\text{strongly negative } (--);\ \text{negative } (-);$$
$$\text{neutral } (N);\ \text{positive } (+); \qquad (2)$$
$$\text{strongly positive } (++)\}.$$

The characteristics of large social systems are determined in the form of statistical distributions. To obtain initial data, we propose a methodology of social system research based on representative sampling. A special questionnaire with a set of direct and indirect questions is used to establish:

– the average number of contacts per individual;

– the level of sociability of the participants;

– the susceptibility indicators of social system's participants, estimated in terms of the set (1), i.e., as the corresponding shares of the total number (ω^L, ω^M, and ω^H);

– the initial distribution of opinions on a given issue (fuzzy assessments from the set (2)) as the corresponding shares ν_0^{--}, ν_0^{-}, ν_0^{N}, ν_0^{+}, and ν_0^{++}.

The social system's participants are represented as separate and interacting elements (agents). Information is disseminated in the system through interpersonal data exchange. By assumption, the behavior of participants obeys the following rules:

- Participants with a high level of sociability and a pronounced (positive or negative) attitude toward the information received share this information.
- During information exchange, the participants with medium and high susceptibility change their opinions under the impact of others when receiving emotionally colored feedback; in the latter case, the opinion can change dramatically (e.g., from positive to negative or from strongly negative to neutral).

Information interaction starts at step $t = 0$, when information is introduced into the social system through the so-called initiating set (a finite number of its representatives who have received information from the primary source). The goal of modeling is to calculate the share of participants who have received information and the distribution of their opinions at each step $t = t + 1$. One step is the time required for the single implementation of all communicative links between the participants.

Social information dissemination processes can be studied using a multidisciplinary approach (mathematics, sociology, and psychology of communication) to solve a wide range of related problems. In particular, the coverage of the target audience by various alarm means in an emergency was determined in the earlier paper [28]. The goal was to develop a methodology for selecting and justifying the parameters of disseminated information blocks as well as their structure and content considering the socio-psychological features of their perception. The emotional component of the social opinion about the threats of emergencies and their consequences was quite obvious. When disseminating information, the understanding of its utility and necessity in several categories ("harmful," "neutral," and "useful") came to the fore.

The number K_{t+1} of informed participants at each interaction step $(t + 1)$ was analytically expressed through the following parameters:

- L, the size of the initiating set;
- $\bar{b}$, the connectivity coefficient of the social system (the average number of links between its members);
- K_t, the number of informed participants at the previous step t;
- q_t, the share of participants willing to share the information received at step t.

The value q_t depends on the share of participants with a high level of sociability and on the relevance of the disseminated information at step t. Sociability is a permanent property of the social system's participants; hence, the level of sociability can be supposed constant. As a rule, the relevance of information decreases over time; for each step, it is calculated using a special relevance decline coefficient and the forecasted information life cycle. However, according to practical evidence, an acute and significant problem retains its relevance in the information space for a long time in emergencies threatening the life and health of people.

Thus, the number of informed agents is given by [28]

$$K_{t+1} = K_t + q_t\left(\frac{N-K_t}{N}\right)\left(K_t^{++}+K_t^{--}\right)\bar{b},$$

where N is the total size of the social system (the region's population). The coefficient $(N-K_t)/N$ reflects the share of uninformed participants at the previous step whereas K_t^{++} and K_t^{--} are the numbers of participants with pronounced attitudes (strongly positive and strongly negative, respectively).

Under the mass spread of infectious disease and the introduction of epidemic safety measures, it is interesting to forecast and study the spectrum of opinions at each information exchange step. The goal is to regulate social opinions by providing a correct understanding of the current epidemic situation and counteracting the dissemination of destabilizing and harmful information.

The distribution of opinions is described by the number of information exchange participants in each category from the set (2). The starting values $(K_0^{--}, K_0^{-}, K_0^{N}, K_0^{+}, K_0^{++})$ are determined within the initiating set L according to the given initial distribution of opinions.

The number of participants with strongly negative attitudes towards information at each step evolves as follows [28]:

$$K_{t+1}^{--} = K_t^{--} + \left(K_{t+1}-K_t\right)\Big[\nu_0^{--} - \nu_0^{--}\times\left(\omega^M+\omega^H\right)\times\left(\frac{K_t^{++}}{K_t^{++}+K_t^{--}}\right) + \nu_0^{-}\left(\omega^M+\omega^H\right)\left(\frac{K_t^{--}}{K_t^{++}+K_t^{--}}\right) + \nu_0^{N}\,\omega^H\left(\frac{K_t^{--}}{K_t^{++}+K_t^{--}}\right)\Big].$$

The factors $\frac{K_t^{++}}{K_t^{++}+K_t^{--}}$ and $\frac{K_t^{--}}{K_t^{++}+K_t^{--}}$ represent the shares of information exchange participants sharing strongly positive and strongly negative opinions, respectively. The formula clearly reflects how moderately and highly susceptible participants are influenced and change their opinions in a certain direction.

The number of participants with a strongly positive opinion is calculated by analogy [28]:

$$K_{t+1}^{++} = K_t^{++} + \left(K_{t+1}-K_t\right)\Big[\nu_0^{++} - \nu_0^{++}\times\left(\omega^M+\omega^H\right)\left(\frac{K_t^{--}}{K_t^{++}+K_t^{--}}\right) + \nu_0^{+}\left(\omega^{\mathrm{M}}+\omega^{\mathrm{H}}\right)\times\left(\frac{K_t^{++}}{K_t^{++}+K_t^{--}}\right) + \nu_0^{N}\,\omega^H\left(\frac{K_t^{++}}{K_t^{++}+K_t^{--}}\right)\Big].$$

The number of participants with a positive opinion is given by

$$K_{t+1}^{+} = K_t^{+} + \left(K_{t+1}-K_t\right)\Big[\nu_0^{+} - \nu_0^{+}\left(\omega^M+\omega^H\right)\times\left(\frac{K_t^{++}}{K_t^{++}+K_t^{--}}\right) - \nu_0^{+}\left(\omega^M+\omega^H\right)\left(\frac{K_t^{--}}{K_t^{++}+K_t^{--}}\right) + \nu_0^{++}\,\omega^M\left(\frac{K^{--}}{K^{++}+K^{--}}\right) + \nu_0^{-}\,\omega^H\left(\frac{K_t^{++}}{K_t^{++}+K_t^{--}}\right) + \nu_0^{N}\,\omega^M\left(\frac{K_t^{++}}{K_t^{++}+K_t^{--}}\right)\Big].$$

Finally, the number of participants with neutral attitude is calculated as

$$K_{t+1}^{N} = K_t^{N} + \left(K_{t+1}-K_t\right)\Big[\nu_0^{N} - \nu_0^{N}\left(\omega^M+\omega^H\right) + \nu_0^{-}\,\omega^M\left(\frac{K_t^{++}}{K_t^{++}+K_t^{--}}\right) + \nu_0^{+}\,\omega^M\left(\frac{K_t^{--}}{K_t^{++}+K_t^{--}}\right) + \nu_0^{++}\,\omega^H\left(\frac{K_t^{--}}{K_t^{++}+K_t^{--}}\right) + \nu_0^{--}\,\omega^H\left(\frac{K_t^{++}}{K_t^{++}+K_t^{--}}\right).$$

4. GENERAL PATTERNS OF INFORMATION INTERACTION: ANALYSIS AND APPLICATION IN CONTROL PROBLEMS

Figure 1 shows typical graphs obtained by computer simulations for the following model dataset:

▪ The social system is composed of 350 000 participants.

▪ The initiating set is 12%.

▪ The average number of links varies from 1 to 5 for 82% of participants and from 6 to 15 for 18% of participants.

▪ The initial distribution of opinions on the issue (the theme of the information disseminated) is specified by $\nu_0^{++} = 0.18$ (strongly positive), $\nu_0^{+} = 0.35$ (positive), $\nu_0^{N} = 0.2$ (neutral), $\nu_0^{-} = 0.17$ (negative), and $\nu_0^{--} = 0.1$ (strongly negative).

▪ The willingness to disseminate the information received is $q_0 = 0.3$.

▪ The levels of susceptibility are $\omega^L = 0.44$ (low), $\omega^M = 0.4$ (medium), and $\omega^H = 0.16$ (high).

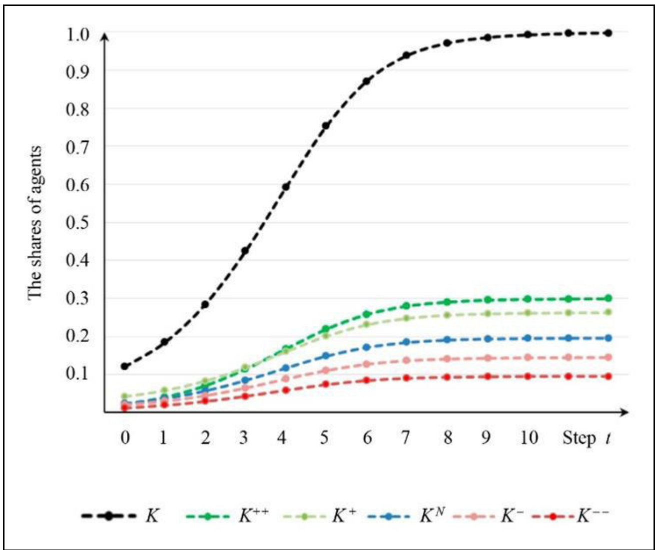

Fig. 1. The growing shares of informed participants and the distribution of their opinions.

According to the experiments, the growing shares of participants in each category represent a sigmoid, i.e., a smooth monotonically increasing nonlinear function with an *S*-shape. The growth is usually smooth. Sharp jumps in the shares of informed participants are possible only for very large (limit) levels of susceptibility and sociability; as surveys show, they are not typical for real social systems.

The proposed model can be used to assess information exchange dynamics under given conditions and forecast social opinions on different issues such as:

▪ the introduction of restrictive measures in a pandemic,

▪ possible shortages of medicines or protective equipment,

▪ the load on the healthcare system,

▪ vaccine development and application,

▪ decreased economic activities (trade, tourism, and culture),

▪ protests denying the danger of the virus and the need for epidemic safety measures.

The resulting forecast based on artificial data simulations can be assessed as unsatisfactory. In this case, the control problem is to choose a control action for the system that will change the level of awareness and the opinion distribution vector to the targets with minimum costs.

Control actions consist in varying the parameters of the social system, e.g., by selecting an appropriate information dissemination channel (i.e., the size of the initiating set) and adjusting the information block's content (i.e., the initial distribution of opinions on issues) or its format (i.e., the willingness to disseminate this block). Also, the formation and detailed planning of available control actions are in the competence of sociologists and psychologists specializing in susceptibility.

The population's reaction is forecasted to make justified managerial decisions and stabilize public sentiment. It is important to plan and disseminate complete and convincing information through trusted media channels for the population to realize the epidemical evolution properly.

The following actions can be performed to counteract the spread of destructive information and avoid destabilization: decreasing the level of susceptibility to destructive information (by reducing trust in its sources) or decreasing the willingness to disseminate it (the loss of relevance compared to other information resources).

Figure 2 shows information dissemination dynamics under the following corrections to the experimental data:

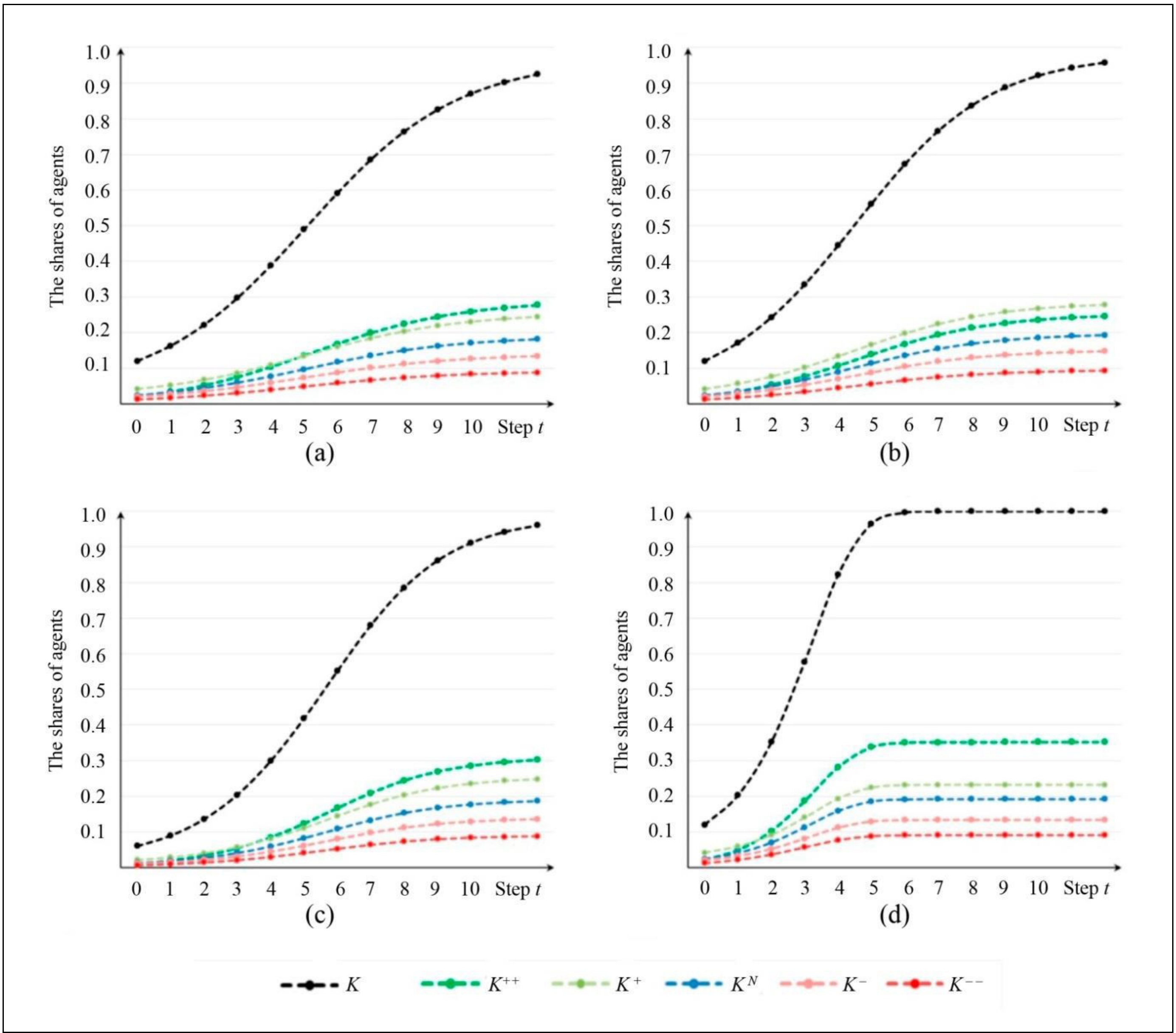

Fig. 2. Information dissemination dynamics under an intervention strategy.

a) decreasing the willingness to disseminate the information received: $q_0 = 0.2$;

b) decreasing the level of susceptibility: $\omega^L = 0.6$ (low), $\omega^M = 0.32$ (medium), and $\omega^H = 0.08$ (high);c) reducing the size of the initiating set: $L = 6\%$;

d) increasing the parameter values: $q_0 = 0.45$, $\omega^L = 0.24$, $\omega^M = 0.53$, and $\omega^H = 0.23$.

Compared to Fig. 1, the graphs in Figs. 2a–2c (Fig. 2d) show a decrease (an increase, respectively) in the speed of information dissemination over the social system.

Figure 3 presents an example with no opinion polarization on the issue and a predominantly neutral attitude at the time of launching the information block $\left(\nu_0^{++} = 0.08,\quad \nu_0^{+} = 0.14,\quad \nu_0^{N} = 0.61,\quad \nu_0^{-} = 0.12,\ \nu_0^{--} = 0.05\right)$:

a) for the levels of susceptibility $\omega^L = 0.2$, $\omega^M = 0.6$, and $\omega^H = 0.2$;

b) for the decreased levels of susceptibility $\omega^L = 0.51$, $\omega^M = 0.42$, and $\omega^H = 0.07$.

The model demonstrates that in the numerical example under consideration, timely measures reducing the level of trust in the source affect the distribution of positive and negative opinions while maintaining an overall neutral information background.

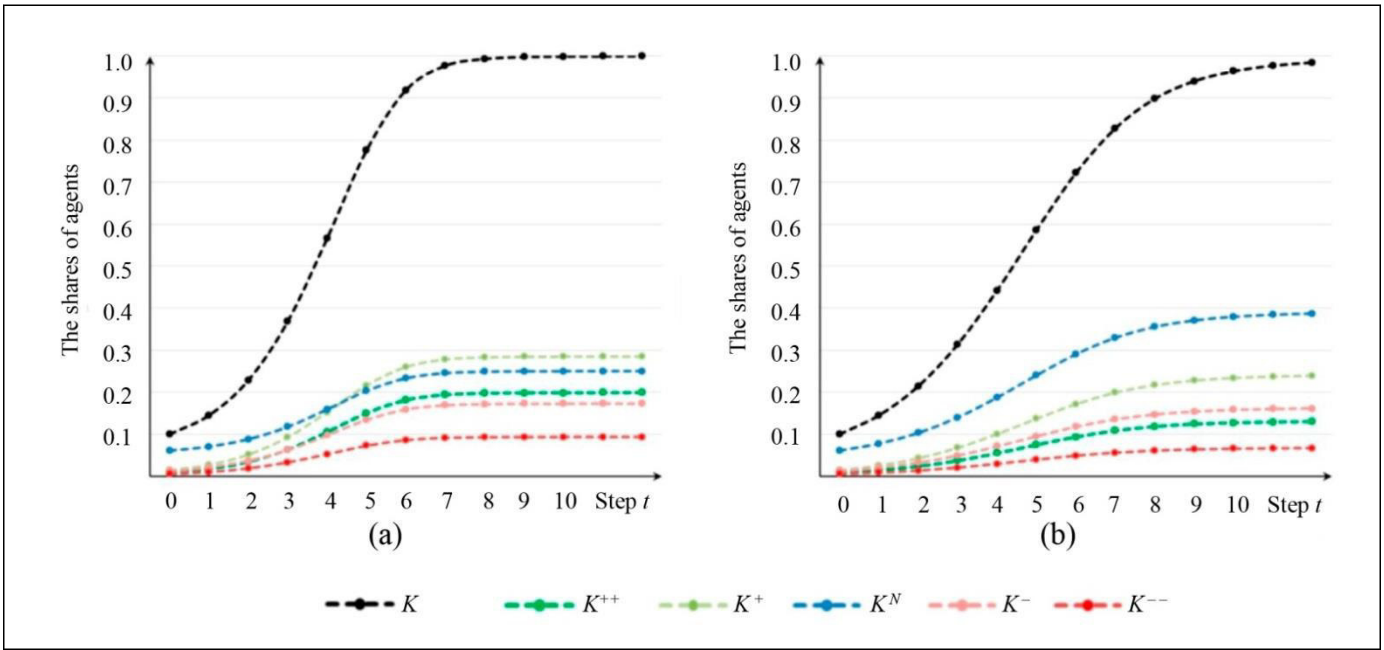

Fig. 3. The distribution of opinions under an intervention strategy.

CONCLUSIONS

Like any emergency, an epidemic generates many negative consequences, influencing almost all sides of human life and being a source of severe psychological stress.

In addition to purely medical aspects, social processes directly or indirectly affect the spread of the disease. Situation analysis and modeling from this point of view assist in organizing competent preventive and psychological corrective measures among the population, reducing the level of stress, and stabilizing social relations.

The mathematical model considered in this paper can be used to forecast the dynamics of social opinions on the applied measures of epidemic safety and to develop managerial decisions toward the adoption of these measures by the population. A numerical example of the distribution of opinions under given characteristics of the social system has been provided; the possibility of changing the dynamics of information dissemination by an intervention strategy has been demonstrated.

In simulation experiments, it is of interest to study several organizational aspects of social systems as follows: clustering of the interaction network of participants and the probability of localization of disseminated information in one cluster; the motivation of participants, incentives, and restrictions in information dissemination; the total susceptibility threshold of a participant simultaneously subjected to several information blocks and sources (as a result, the participant may neglect some of them and not react to the control action). Also, it is necessary to analyze real data from online social media. Further research in these directions seems topical and promising.

REFERENCES

1. Asatryan, M.N., Gerasimuk, E.R., Logunov, D.Yu., et al., Predicting the Dynamics of Covid-19 Incidence and Planning Preventive Vaccination Measures for Moscow Population Based on Mathematical Modeling, *Journal of Microbiology, Epidemiology and Immunobiology*, 2020, vol. 97 (4), pp. 289–302. (In Russian)]
2. Vytla, V., Ramakuri, S.K., Peddi, A., et al., Mathematical Models for Predicting Covid-19 Pandemic: a Review, *Journal of Physics: Conference Series*, 2021, vol. 1797. DOI: 10.1088/1742-6596/1797/1/012009.
3. Rahman, A., Kuddus, M.A., Ip, R.H.L., and Bewong, M., A Review of COVID-19 Modelling Strategies in Three Countries to Develop a Research Framework for Regional Areas, *Viruses*, 2021, vol. 13, art. no. 2185. URL: https://doi.org/10.3390/v13112185.
4. Osipov, V.Yu., Kuleshov, S.V., Zaytseva, A.A., and Aksenov, A.Yu., Approach for the COVID-19 Epidemic Source Localization in Russia Based on Mathematical Modeling, *Informatics and Automation*, 2021, vol. 20, no. 5, pp. 1065–1089. (In Russian.)

5. Krivorot'ko, O.I., Kabanikhin, S.I., Zyat'kov, N.Yu., et al., Mathematical Modeling and Forecasting of COVID-19 in Moscow and Novosibirsk Region, *Numerical Analysis and Applications*, 2020, vol. 13, no. 4, pp. 332–348.
6. Martyanova, A.Ye. and Azhmukhamedov, I.M., SEIRD Model Describing the Dynamics of the Spread Viral Infections Considering the Appearance of New Strains, *Caspian Journal: Control and High Technologies*, 2022, no. 4(60), pp. 36–43. (In Russian.)
7. Zakharov, V.V. and Balykina, Yu.E., Balance Model of COVID-19 Epidemic Based on Percentage Growth Rate, *Informatics and Automation*, 2021, vol. 20, no. 5, pp. 1034–1064. (In Russian.)
8. Makarov, V.L., Bakhtizin, A.R., Sushko, E.D., and Ageeva, A.F., COVID-19 Epidemic Modeling – Advantages of an Agent-Based Approach, *Economic and Social Changes: Facts, Trends, Forecast*, 2020, vol. 13, no. 4, pp. 58–73. (In Russian.)
9. Sokolovsky, V.L., Furman, G.B., Polyanskaya, D.A., and Furman, E.G., Spatio-Temporal Modeling of COVID-19 Epidemic, *Health Risk Analysis*, 2021, no. 1, pp. 23–37.
10. Uttrani, S., Nanta, B., Sharma, N., and Dutt, V., Modeling the Impact of the COVID-19 Pandemic and Socioeconomic Factors on Global Mobility and Its Effects on Mental Health, in *Artificial Intelligence, Machine Learning, and Mental Health in Pandemics: a Computational Approach*, 1st ed., Chapter 8, Amsterdam: Elsevier, 2022, pp. 189–208. DOI:10.1016/B978-0-323-91196-2.00012-0.
11. Sorokin, M.Yu., Kasyanov, E.D., Rukavishnikov, G.V., et al., Psychological Reactions of the Population as a Factor of Adaptation to the COVID-19 Pandemic, *V.M. Bekhterev Review of Psychiatry and Medical Psychology*, 2020, no. 2, pp. 87–94. (In Russian.)
12. Belyakov, N.A., Bagnenko, S.F., Rassokhin, V.V., et al., *Evolyutsiya pandemii COVID-19* (The Evolution of COVID-19 Pandemic), St. Petersburg: Baltic Medical Education Center, 2021. (In Russian.)
13. Simoy, M.I. and Aparicio, J.P., Socially Structured Model for COVID-19 Pandemic: Design and Evaluation of Control Measures, *Computational and Applied Mathematics*, 2022, vol. 41, no. 1. DOI: https://doi.org/10.1007/s40314-021-01705-1.
14. Nowzari, C., Preciado, V.M., and Pappas, G.J., Analysis and Control of Epidemics: a Survey of Spreading Processes on Complex Networks, *IEEE Control Systems Magazine*, 2016, vol. 36, pp. 26–46. DOI: 10.1109/MCS.2015.2495000.
15. Merler, S., Ajelli, M., Fumanelli, L., et al., Spatiotemporal Spread of the 2014 Outbreak of Ebola Virus Disease in Liberia and the Effectiveness of Non-pharmaceutical Interventions: a Computational Modelling Analysis, *The Lancet Infectious Diseases*, 2015, vol. 15, no. 2, pp. 204–211. DOI: 10.1016/S1473-3099(14)71074-6.
16. Gubar, E., Taynitskiy, V., Fedyanin, D., and Petrov, I., Hierarchical Epidemic Model on Structured Population: Diffusion Patterns and Control Policies, *Computation*, 2022, vol. 10, no. 31. DOI: 10.3390/computation10020031.
17. Fenichel, E.P., Economic Considerations for Social Distancing and Behavioral Based Policies During an Epidemic, *J. Health Econ.*, 2013, vol. 32, no. 2, pp. 440–451. DOI:10.1016/j.jhealeco.2013.01.002.
18. Flache, A., Mäs, M., Feliciani, T., et al., Models of Social Influence: towards the Next Frontiers, *Journal of Artificial Societies and Social Simulation*, 2017, vol. 20, no. 4. URL: http://jasss.soc.surrey.ac.uk/20/4/2.html.
19. Gubanov, D.A., Petrov, I.V., and Chkhartishvili, A.G., Multidimensional Model of Opinion Dynamics in Social Networks: Polarization Indices, *Control Sciences*, 2020, no. 3, pp. 26–33. (In Russian.)
20. Zhang, L., Li, K., and Liu, J., An Information Diffusion Model Based on Explosion Shock Wave Theory on Online Social Networks, *Appl. Sci.*, 2021, no. 11. URL: https://doi.org/10.3390/app11219996.
21. Gubanov, D.A., A Study of a Complex Model of Opinion Dynamics in Social Networks, *Journal of Physics: Conference Series*, 2021, vol. 1740, pp. 1–6. URL: https://iopscience.iop.org/article/10.1088/1742-6596/1740/1/012040/pdf.
22. Gubanov, D.A., Kozitsin, I.V., and Chkhartishvili, A.G., Face Mask Perception During the COVID-19 Pandemic: an Observational Study of Russian Online Social Network VKontakte, *Advances in Systems Science and Applications*, 2021, vol. 21, no. 3, pp. 91–100.
23. Milov, O., Yevseiev, S., Milevskyi, S., et al., Critical Points of Information Influence in Social Networks, *Proceedings of the 3rd International Scientific and Practical Conference "Information Security and Information Technologies,"* Kirovograd, 2021. URL: https://ceur-ws.org/Vol-3200/paper18.pdf.
24. Al-Oraiqat, A., Ulichev, O.S., Meleshko, Ye., et al., Modeling Strategies for Information Influence Dissemination in Social Networks, *Journal of Ambient Intelligence and Humanized Computing*, 2022, no. 13(91), pp. 2463–2477. DOI:10.1007/s12652-021-03364-w.
25. Perra, N. and Rocha, L.E., Modelling Opinion Dynamics in the Age of Algorithmic Personalization, *Scientific Reports*, 2019, vol. 9, art. no. 7261.
26. Parsegov, S.E., Proskurnikov, A.V., Tempo, R., and Friedkin. N.E., Novel Multidimensional Models of Opinion Dynamics in Social Networks, *IEEE Transactions on Automatic Control*, 2017, vol. 62, no. 5, pp. 2270–2285.
27. Kozitsin, I.V., Formal Models of Opinion Formation and Their Application to Real Data: Evidence from Online Social Networks, *The Journal of Mathematical Sociology*, 2020, vol. 46, pp. 120–147. DOI: 10.1080/0022250X.2020.1835894.
28. Azhmukhamedov, I.M., Machueva, D.A., and Alisultanova, E.D., Mathematical Modeling of Information Management of Social Systems in Emergencies, *Advances in Economics, Business and Management Research*, 2020, vol. 156, pp. 695–701.

This paper was recommended for publication
by RAS Academician D.A. Novikov,
a member of the Editorial Board.

Received December 12, 2022,
and revised September 19, 2023.
Accepted September 19, 2023.

Author information

Azhmukhamedov, Iskandar Maratovich. Dr. Sci. (Eng.), Astrakhan State University, Astrakhan, Russia
✉ aim_agtu@mail.ru
ORCID iD: https://orcid.org/0000-0001-9058-123X

Machueva, Dina Aluevna. Cand. Sci. (Eng.), Grozny State Oil Technical University, Grozny, Russia
✉ ladyd_7@mail.ru
ORCID iD: https://orcid.org/0000-0001-6827-8998

Cite this paper
Azhmukhamedov, I.M., Machueva, D.A., Modeling Social Attitude to Introducing Epidemic Safety Measures in a Pandemic. *Control Sciences* **5,** 56–64 (2023). http://doi.org/10.25728/cs.2023.5.5

Translated into English by *Alexander Yu. Mazurov*,
Cand. Sci. (Phys.–Math.),
Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
✉ alexander.mazurov08@gmail.com

DOI: http://doi.org/10.25728/cs.2023.5.6

EXPANDING THE FUNCTIONALITY OF AN APPLIED GEOGRAPHIC INFORMATION SYSTEM FOR MODELING SEARCH CORRELATION-EXTREME NAVIGATION SYSTEMS

A.I. Alchinov[1] and I.N. Gorokhovsky[2]

[1]Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
[2]Research Center of Topographic and Navigational Support, Central Research Institute No. 27, Moscow, Russia

✉ alchinov46@mail.ru, ✉ gin_box@mail.ru

Abstract. This paper further develops the concept of an applied geographic information system (AGIS) for modeling search correlation-extreme navigation systems (CENSs) intended to control moving objects. The possibility of using parallel, distributed, and cloud computing for modeling CENSs is investigated. In modern conditions, it is necessary to diagnose the operation of CENSs under stress exposure on their shooting systems. The stress exposure parameters are modeled by accessing specialized databases containing the characteristics of terrain objects in different electromagnetic radiation wavelength ranges. As a rule, such characteristics are unavailable in geographic information systems (GISs) and cloud environments. It is demonstrated that CENSs should be diagnosed by modeling the shooting system using cloud GISs. The issues of parallel computing for pattern recognition tasks are considered. The peculiarities of the parallel structure of CENS search algorithms are revealed. When implementing these algorithms in parallel computing systems, proper consideration of the peculiarities allows utilizing their advantages to the highest degree.

Keywords: cloud computing, cloud service, parallel computing, computing systems, pattern recognition, cloud geographic information system, search correlation-extreme navigation system.

INTRODUCTION

Onboard search correlation-extreme navigation systems (CENSs) for various-purpose unmanned vehicles (marine, aerospace, and ground) determine the parameters of their motion by checking hypotheses about the values of these parameters: CENSs match the current terrain sector image received by the onboard shooting system in the next autonomous navigation session in a given area with fragments of a reference image of this area. The reference images are prepared in advance and are stored in the memory of the onboard computer. When searching for a reference image fragment close by content to the terrain sector image (in the sense of a closeness function in the onboard algorithm), a regular shift grid of the frame of the appropriate size and orientation is used to select the next reference image fragment. The hypotheses that the sought parameters have values equal to those at the grid nodes are checked. The hypothesis for which the closeness function achieves maximum is accepted. Global (exhaustive) and local (gradient) optimization schemes are applied. According to the authors' publications [1, 2], autonomous navigation in modern conditions requires developing more general methods for checking hypotheses about the values of the motion parameters of unmanned vehicles than those described above. In this paper, such methods will be called overview and comparative methods. For the technical systems implementing these methods, we will retain the above name and designation established historically.

The topicality of the accelerated development of CENSs at the present stage was justified in [1, 2]. As was shown therein, a promising R&D line is the creation of an applied geographic information system (AGIS) for modeling CENSs. This system should be equipped with the means of assembling computer models of a wide range of CENS variants and draft technologies to adjust their operating sessions in specified areas from off-the-shelf software components for conducting computational experiments in order to assess the effectiveness and stability against stress exposures.

The authors proposed a mathematical model [1] to justify the composition of such software components as well as their functionality and spatial data requirements. In Section 1 below, this mathematical model is applied to identify (and describe in terms of related problem domains) some analogs of software components and spatial data of AGIS CENSs. Under appropriate access to such components and data, the implementation costs of AGIS CENSs can be minimized. Also, the general parallel structure of CENS search algorithms is given a compact description. This description can be used when implementing AGIS CENSs based on parallel computing systems.

Sections 2 and 3 are devoted to the possibility and options for accessing analogs of AGIS CENSs components in related problem domains.

Many images in different electromagnetic radiation (EMR) wavelength ranges are used in the development of models and layouts of search CENSs and their adjustment for operation. Image processing on computing systems involves large amounts of data; much processor time is spent on simulating CENSs [1–3]. The performance of computing systems should be improved using parallel computing in modern cloud environments (cloud GISs), the so-called distributed GISs with decentralized control in the parallel execution of simulation tasks [4, 5].

Cloud environments permanently store large amounts of data (remote sensing images of the earth), e.g., a hypercube of information that will be temporarily cached on the user's side. The user also gets quick access to the data, while saving time and resources. The matter concerns distributed GISs belonging to one or several organizations dealing with CENSs. Software tools, the structures of storage devices and the interface of AGIS CENSs, and all the methods (pattern recognition, scene analysis, clustering, and training of neural networks) can be implemented using network services with fast search of necessary information [3, 6, 7].

Currently, these tasks are performed using the well-known classical approaches on computers that are linked to server databases within one or several organizations. Each group of tasks is performed in dedicated applications, and it is possible to use different approaches when decomposing the tasks. The methods and means to implement parallelism depend on the level where they should be provided. The capabilities and means to implement parallel computing depend on the level of commands, data flows, and tasks.

The authors outlined the further development of search CENSs towards new design principles for onboard algorithms, their intellectualization and self-organization, the application of new types of shooting systems and their combination, and the implementation of parallel CENS algorithms, including the capabilities of cloud technologies.

The accelerated development of CENSs has become topical due to the emergence of parallel computing-oriented processors and the development of software tools for training neural networks on big data in cloud computing environments [1, 2].

In view of these conditions, R&D works have been expanded using a computational complex that provides all the necessary means for, first, assembling models of a wide range of CENSs and draft technologies to adjust their operation in specified areas from off-the-shelf software components via a special interface to databases and, second, conducting experiments.

The functionality to model search CENSs and draft technologies for their adjustment can be implemented in a general-purpose GIS using universal tools for handling geospatial information. In Sections 2 and 3, this functionality is expanded based on the analysis of cloud technologies and parallel computing. Also, we investigate the capabilities of cloud technologies and parallel computing, the peculiarities of their application when modeling advanced CENSs, and their possible implementation based on data processing means and technologies.

A computing system supporting parallel calculations achieves maximum efficiency at the level ensuring system parallelism with the decomposition of tasks that can be performed simultaneously [1, 2]. The approaches to calculations in modern computing systems are analyzed, and the data necessary for choosing a computing system for complex tasks are obtained. We decompose one of the main tasks of AGIS CENSs: modeling CENS operation and determining the peculiarities of the parallel structure of the corresponding algorithms.

This study expands the functionality of the existing general-purpose GISs to model search CENSs effectively. To achieve this goal, we further develop the mathematical model of search CENSs. Based on this model, we decompose the tasks of AGIS CENSs, including those of modeling stress exposures on their imaging systems. We analyze general-purpose cloud technologies and existing cloud GISs designed for processing geospatial information in terms of their application to the tasks mentioned above. The results presented below can be used when implementing AGIS CENSs.

1. THE PARALLEL STRUCTURE OF CENS MODELING ALGORITHMS. SOME ANALOGS OF SOFTWARE COMPONENTS OF AGIS CENSS IN RELATED PROBLEM DOMAINS

The starting points are as follows.

• The maximum efficiency of the computing system selected for implementing AGIS CENSs can be achieved if its characteristics are chosen considering the peculiarities of the initial data and computer algorithms to simulate CENS adjustment processes for autonomous orientation using overview and comparative methods and the operation processes of the adjusted CENS under various disturbances, including stress factors. Such processes are simulated by means of contemporary computing systems at the limit of their performance. Besides the available memory and the speed of processors and exchange devices between them, the key role is played by their specialization and the possibility of parallelizing the subtasks of general tasks during simulation [4]. The central general task in AGIS CENSs is to simulate the operation of CENSs adjusted for autonomous orientation. The peculiarities of the parallel structure of algorithms to perform this general task are considered below. The peculiarities of simulating CENS adjustment processes and image synthesis processes for different shooting systems used in CENS will be addressed in subsequent publications of the authors.

• In order to determine the specialization of computing system processors and organize their parallel operation for the central general task of AGIS CENSs, it is necessary to decompose this task into subtasks and to distinguish those that must be performed sequentially (the output of a previous subtask is the input for the subsequent one) and those that can be performed in parallel.

The decomposition can be performed using the mathematical models proposed in [2, 3]. As in these papers, we consider search CENSs in which the shooting system captures a scene image S on a terrain section, and the onboard algorithm refines the planned coordinates $d=(X,Y)$ of the carrier at the shooting instant. Let M be the set of all possible images S coming from the shooting system to the input of the onboard algorithm of the CENS at the shooting instant. Then the search CENS adjusted for autonomous orientation can be treated as a calculator of the values of a function $\hat{f}(S): M \to \hat{D}$ fixed during the adjustment process for any "value" of the "variable" $S \in M$ coming from the shooting system. In other words, upon receiving an image S, the CENS will calculate and output the value of the function $\hat{f}(S)=\hat{d} \in \hat{D}$ (its response). The finite set $\hat{D}$ combines all variants of the computer's responses to the question about the carrier's location at the instant of receiving the image S. In [1, 2], a general analytical expression was derived for generalized step functions defined on hierarchical partitions of the set M into disjoint classes, in which a different prior image transformation can be applied for each partition level. In this paper, we consider $r = 2$ levels of hierarchical partitioning into classes and subclasses. This number is sufficient for decomposing the general task and revealing the peculiarities of parallelism. In addition, with a high degree of confidence, this variant will attract major interest in the practical use of CENSs.

In this case, we have

$$M=\bigcup_{i=1}^{l} K_i, \text{ where } K_m \cap K_n = \varnothing \ \ \forall m, n \in [1, l], \ m \neq n,$$

$$K_i=\bigcup_{j=1}^{l_i} K_{ij}, \text{ where } K_{im} \cap K_{in} = \varnothing$$

$$\forall i = 1, \ldots, l \text{ and } m, n \in [1, l_i], \ m \neq n.$$

Then the function $\hat{f}(S)$ can be written in the generalized vector form

$$\hat{f}(S)=\Big\langle \chi(S), \Big(\big\langle \chi_1(S), \hat{\mathbf{d}}_1 \big\rangle, \big\langle \chi_2(S), \hat{\mathbf{d}}_2 \big\rangle, \ldots, \big\langle \chi_l(S), \hat{\mathbf{d}}_l \big\rangle \Big) \Big\rangle \quad (1)$$

with the following notations: angular brackets indicate the inner product of two vectors; $\chi(S)$ = ($\chi_1(S)$, $\chi_2(S)$,..., $\chi_l(S)$); $\chi_i(S)$ is the characteristic function of class K_i, $i=1,\ldots,l$; $\chi_i(S) = (\chi_{i1}(S),\ \chi_{i2}(S),\ \ldots,\ \chi_{il_i}(S))$; $\langle \chi_i(S), \hat{\mathbf{d}}_i \rangle = \sum_{j=1}^{l_i} \chi_{ij}(S)\hat{d}_{ij}$; $\chi_{ij}(S)$ is the characteristic function of class K_{ij}, $i=1,\ldots,l$, $j=1,\ldots,l_i$; finally, $\hat{\mathbf{d}}_i = (\hat{d}_{i1},\ \hat{d}_{i2},\ldots,\ \hat{d}_{il_i})$.

At the first and second partition levels, the input image S possibly undergoes preliminary transformations $\pi(S)$ and $\pi_i(S)$ of respectively. The algorithms for calculating the values of the vector functions $\chi(\pi(S))$ and $\chi_i(\pi_i(S))$ determine the class of the image S. Therefore, by the fundamental theorem on representing any recognition algorithm through the sequential execution of a recognizing operator and a decision rule, we obtain [8]

$$\chi(S) = \mathbf{C}\,(\mathbf{B}(\pi(S))),$$

where $\mathbf{B}(\pi(S)) = (b_1(\pi(S)),\ b_2(\pi(S)),\ldots,\ b_l(\pi(S)))$; $b_i(\pi(S))$ is a numerical measure of the closeness of the image to the class $K_i \in M$; $\mathbf{C}\,(b_1(\pi(S)),\ b_2(\pi(S)),\ \ldots,\ b_l(\pi(S))) = (c_1,\ c_2,\ldots,\ c_l)$, where $c_i \in \{0, 1\}$, $i=1,\ldots,l$; $\chi_i(S) = \mathbf{C}_i(\mathbf{B}_i(\pi(S)))$, $\mathbf{B}_i$, and $\mathbf{C}_i$ are described by analogy.

Then the function $\hat{f}(S)$ (1) of the CENS computer takes the form

$$\hat{f}(S) = \Big\langle \mathbf{C}\big(\mathbf{B}(\pi(S))\big), \Big(\big\langle \mathbf{C}_1\big(\mathbf{B}_1(\pi_1(S))\big), \hat{\mathbf{d}}_1 \big\rangle, \big\langle \mathbf{C}_2\big(\mathbf{B}_2(\pi_2(S))\big), \hat{\mathbf{d}}_2 \big\rangle, \ldots, \big\langle \mathbf{C}_l\big(\mathbf{B}_l(\pi_l(S))\big), \hat{\mathbf{d}}_l \big\rangle\Big)\Big\rangle. \tag{2}$$

According to (2), the general task includes the subtasks of calculating the "values" of the preliminary transformations involved (hereinafter called the subtasks of type π). The results of performing a subtask of type π are the initial data for the corresponding subtask of calculating the vector of numerical measures of closeness to the partition classes $M = \bigcup_{i=1}^{l} K_i$ (hereinafter called the subtask of type *B*). The results of performing a subtask of type *B* are the initial data for the corresponding subtask of type *C*, which determines the class of the image by analyzing the vector of closeness measures. After deciding that the image belongs to some class of the first partition level, the procedure is repeated for the partition of this class with performing the subtasks of types π, *B*, and *C*. The process is completed by extracting from the computer's memory the value $\hat{d}_{ij}$ (the partition class selected by the decision rule of the second partition level). The subtasks of type *C* are performed only after the corresponding subtask of type *B*, which can be performed only after the subtask of type π. The vector "coordinates" in this process can be computed in parallel; the inner products, only sequentially as soon as the operand values are available. This is the peculiarity of parallelism when modeling CENS operation. According to the classification given in [5], the corresponding computing system has parallelism at the level of threads. In this case, the acceleration of program execution due to parallelizing its instructions on the set of computers is limited by the time required for executing its sequential instructions.

Note that this mathematical model allows implementing AGIS CENSs with the maximum use of off-the-shelf computer components from the problem domains intended for performing the tasks of types π and $\chi(S) = \mathbf{C}(\mathbf{B}(\pi(S)))$ and collecting and generalizing initial data to adjust CENSs for autonomous orientation by overview and comparative methods in a given area.

Based on the mathematical model, we also identify components of AGIS CENSs that have analogs in related problem domains. The cost of implementing AGIS CENSs can be minimized under the required options of access to such components. The connections between the elements of the mathematical model and the software components from related problem domains implementing them are presented in the table below.

The next sections of the paper are devoted to analyzing the capabilities and forms of access to analogs of AGIS CENSs components in related problem domains.

2. ACCESS TO CLOUD TECHNOLOGIES AND CLOUD GISS: THE EXISTING CAPABILITIES AND PROSPECTS

To analyze off-the-shelf components (programs and databases) in the implementation of algorithms for modeling search CENSs, we overviewed cloud technologies and cloud GISs based on scientific research that can be used in GIS applications for modeling search CENSs. Cloud technologies were investigated to identify the existing applications that can serve to model search GISs, particularly the shooting systems of CENSs. No such applications were found.

The mathematical models of AGIS CENSs components (see the table) can be implemented as independent tasks using parallel and distributed computing.

The initial information about orientation conditions in an autonomous navigation area of an unmanned vehicle by the overview and comparative method in the form of relational data (a representative set of *N* "samples") is in storages that can be accessed in an organized way. For distributed GISs, the access is already organized; for cloud GISs, it is necessary to develop appropriate access applications.

The initial information about orientation conditions in an autonomous navigation area of an unmanned vehicle by the overview and comparative method in the form of a computer model of the shooting system includes the following models:

– computer models of the terrain in the areas of autonomous navigation sessions,

– simulation models of the shooting system,

– computer models of stress exposures on the terrain and the images of its sections for different shooting systems.

The first two models have already been implemented in distributed GISs and cloud GISs, in contrast to computer models of stress exposures on shooting systems.

Analogs of AGIS CENSs components in related problem domains

Mathematical models of AGIS CENSs components	Analogs in related problem domains
The initial information about orientation conditions in an autonomous navigation area of an unmanned vehicle by the overview and comparative method in the form of a table containing a representative set of N "samples" $\left(S_j, d_j\right)$ of the function $f(S): M \to D$ that describes these conditions: $$I_0\left\{f(S): M \to D\right\} = \left(S_j, d_j\right),\ d_j \in D,\ j = 1,\dots, N,$$ where D denotes the set of admissible values of the parameter being refined in an autonomous navigation session.	Storages of images of different territories from various moving objects that were obtained using shooting systems similar to those adopted in CENSs.
The initial information about orientation conditions in an autonomous navigation area of an unmanned vehicle by the overview and comparative method in the form of a computer model of the shooting system operating in this area: $$I_0\left\{f(S): M \to D\right\} = \hat{f}^{-1}(d, p),\ d \in D,\ p \in P,$$ where $p \in P$ are the disturbing parameters considered in the shooting system model (one generalized parameter p with the domain P of admissible values). The simulation model of the shooting system should approximate the function $f^{-1}(d)$, $d \in D$ (the inverse of $f(S): M \to D$).	Computer models of shooting systems operating in the areas of autonomous navigation sessions. They have the following components: • computer models of the terrain in such areas; • simulation models of shooting systems that form the images of sections in such areas using a terrain model of the areas similar to real shooting systems; • computer models of stress exposures on the terrain and the images of its sections for different shooting systems.
The onboard computer of CENS: a parametric family of single-valued functions $\left\{\hat{f}(\alpha)(S)\right\}_{\alpha \in A}$, where $\hat{f}(\alpha)(S): M \to \hat{D}$ is a particular function from this family. When CENS is adjusted for autonomous navigation by the overview and comparative method in a given area, the choice of this function is uniquely determined by the value of the generalized parameter $\alpha \in A$. The approximating functions from the parametric family of step functions include components of types π and B; see above.	They have the following components: • type π, all known types of digital image processing operations (filtering, boundary selection, scene description in the areas caught in the frame, segmentation, etc.); • type B, components of all known types of recognition operators within families of pattern recognition algorithms (potential functions, separating surfaces, voting, etc.).

According to the third row of the table, the model of onboard computers includes components of types π and *B*, where π corresponds to known types of digital image processing operations (filtering, boundary selection, scene description for the areas caught in the frame, segmentation, etc.) and *B* to known types of recognition operators within the families of pattern recognition algorithms (potential functions, separating surfaces, voting, etc.). Their implementability in cloud GISs is still under study. The main databases, their structure, and content for possible use in modeling search CENSs were also considered. As was established, terrain images obtained in different EMR wavelength ranges are mostly standardized and are stored in different known formats; their use via cloud services will cause no difficulty. However, the problem lies in the processing of significant amounts of information contained primarily in space images.

As was discovered, cloud environments have no specialized databases with reflecting and absorbing characteristics of terrain. Such characteristics are necessary to form initial navigation reference data in given EMR wavelength ranges and to diagnose CENS operation under stress, when the necessary parameters of stress exposures are obtained considering the parameters of terrain objects.

Currently, there exist the following types of services using cloud computing: *Software as a Service* (SaaS), *Infrastructure as a Service* (IaaS), and *Platform as a Service* (PaaS).

Simple integration solutions involve conventional software tools that define software architecture [5, 9]. Such engineering solutions as containerization, microservice cloud architecture, multi-cloud solutions, and hybrid cloud environments are being intensively developed nowadays. They are directly related to the development of AGIS CENSs, where multiple tasks need to be performed in parallel using cloud technologies for decentralized computing.

At present time, cloud computing is used in different spheres. However, the application of cloud computing to manage the information support of CENSs technologies was not described in the literature [10, 11]. This conclusion follows from the analysis of numerous publications. The explanation is software peculiarities: complexity, high price, and the skill level of those engaged in processing large amounts of data with specified requirements. Databases on different territories are needed for the information support of the CENS technology. They require the performance of large, complex, and expensive works as well as flawless operation of the means of using these databases.

CENS modeling tasks cannot be fully implemented using cloud technologies due to the existing technical and regulatory constraints:

- large amounts of data that are difficult to transmit over networks [12–15];
- limiting requirements for the use of aerospace survey materials, UAVs, or maps and plans of given scale and content.

In recent years, cloud technologies have been widely used to provide Russian government services. Note the National Digital Economy Project and its contribution to market growth and IT development.

The use of cloud technologies is subject to the following constraints: stability and recoverability, security standards, special standards, and compliance with state regulations.

Existing data are mostly unstructured, some are organized using metadata, with a crucial component to unify any data. Virtualization, parallel processing, distributed file systems, and databases used in computers significantly increase the efficiency of big data processing.

The Internet is a distributed computing network that requires powerful computers and, moreover, connection optimizers over WAN channels to improve big data processing. The components of this service need to be distributed among several nodes with the management of a hypervisor that supports the optimal utilization of the IT infrastructure. The structure and interoperable standardized environments of big data are essential for application access. These data are mostly in the form of conventional relational databases. The distribution of applications for parallel computing is the core of any big data solution.

According to the analysis of the available databases, the reflecting and absorbing characteristics are absent for many terrain objects. They are heterogeneous, not systematized, have different formats, and cannot be used to model search CENSs without preliminary processing. The external management of databases is difficult to optimize by software automation. Therefore, automatic data supply, copying, and scaling as well as database access security are problematical. Furthermore, such cloud environments are necessary for specialized organizations that have not yet formed databases of the reflecting and absorbing characteristics of terrain objects in different EMR wavelength ranges. This task requires many years of work.

All these considerations necessitate the use of distributed GISs, e.g., cloud GISs. Such a solution provides access to many types of information that can significantly increase the capabilities of GIS technologies to create new types of thematic GISs. Different

GISs can form their own databases, available to other users as a cloud service. Note the following aspect concerning cloud GISs: by now, GIS capabilities have almost not changed, and only the possibility of virtual work with GISs has been implemented.

The next section describes cloud technologies with parallel computing in CENS modeling tasks.

3. DISTRIBUTED AND CLOUD COMPUTING IN APPLIED GISS FOR MODELING CENSS

Distributed computing to perform tasks in applied GISs for modeling CENSs can be implemented on separate computers united in a parallel computing system. Such calculations can be in one organization or several specialized organizations. The entire package of tasks performed by an applied GIS can be arranged on one computer in a general-purpose GIS in the case of no time constraints. This approach involves a distributed database in the form of a set of relations constituting a single dataset. It can be accessed when performing any task from the set of tasks with simultaneous assignment of tasks to other computers, servers, and cloud services. Also, note parallel databases with a dedicated management system for parallelization of CENS modeling tasks.

The tasks performed in applied GISs in computing systems are represented as separate applications with different task decomposition methods (algorithms, hypotheses, etc.; see below). Nowadays, IT experts use mainly GIS technologies, e.g., in thematic cartography. On the one hand, cloud technologies contribute to IT development; on the other hand, not every platform can be used to process large amounts of data, including complex and bulky software.

Cloud services can represent data repositories for GIS technologies used to model CENS. These are aerospace images of different scales, as well as images obtained from UAVs in different EMR wavelength ranges. Images of appropriate resolution are used for high-accuracy navigation. Image processing requires applications, which are now being developed by some companies. For example, Esri considers cloud technologies one of the directions to develop the ArcGIS platform. A cloud variant of ArcGIS Server 10 is embedded in Amazon's cloud infrastructure. Note that the software is not installed in the cloud environment: the corresponding functions remain in-house. This is a new technology for cameral and field surveying, where a cloud GIS provides the necessary access to data and tools.

Russian GIS products, both multipurpose and distributed, have been developing following the global trends in geoinformatics, i.e., client-server applications and support for Oracle Spatial [16].

The main Russian GISs are as follows: Panorama (Panorama Design Bureau), Talka-GIS (Trapeznikov Institute of Control Sciences, Russian Academy of Sciences), GeoGraf (Mining Exploration Center, the Institute of Geography, Russian Academy of Sciences), PARK (The All-Russian Geological Research Institute (VSEGEI), Moscow Branch), Sinteks ABRIS (TRISOFT LLC), and InGeo (Integro, Center for System Studies, Ufa) [17].

Panorama, the most famous GIS in Russia, also employs cloud technologies on the Panorama @ GeoCloud GIS station with a control system for selecting computing power and time of use. Developed by an Israeli company, this service provides remote work with software products, databases, and the best software solutions in GIS technologies. Currently, this station creates digital maps, processes remote sensing data, performs various measurements, and builds 3D terrain models from cloud databases created, however, by an Israeli company. This fact may be unsuitable for Russian users.

The current interest in cloud technologies does not reduce the interest in conventional GISs. It is possible to talk about distributed GISs as a service (cloud GISs), an additional platform to expand engineering solutions and optimize production costs simultaneously [14]. Hardware and software means provided and distributed through networks is an important line to develop new-generation GISs. Therefore, GIS tools for modeling CENS still remain more attractive. Data retrieval (visualization and loading if necessary) and data transformation according to the established requirements occur when accessing the appropriate services; these functions are implemented in GIS. Cloud service support is possible in terms of providing images, the spectral brightness coefficients of terrain objects in different EMR wavelength ranges. These services are departmental.

As a rule, GIS applications are implemented in one operating system. Simultaneous access to multi-task applications is possible; in this case, task calculations in applications have to be parallelized. When implementing several applications in GIS, it is necessary to speak about multithreading. In particular, this applies to the creation of disturbances for the operation of CENSs (stress exposures on the shooting system), random or purposeful [1, 2]. Assessing the performance of the shooting system of CENSs subjected to stress exposures is complex and multi-cyclic; it requires considerable processor time.

ArcGIS Pro is a multithreaded application. It offers over 70 surveying tools supporting parallel computing. As a rule, when working with ArcGIS Pro, it is possible to process geodetic data in the background mode and create maps and work with other scenes simultaneously [19]. However, no enterprises in Russia currently use this software product with vendor's technical support.

In GIS Panorama (ver. 12, Russia), multithreaded file processing is implemented. As a result, the performance of data processing is increased almost linearly (by the number of processor cores) based on MAPAPI interface functions adapted to multithreading [19].

Digital photogrammetric station (DPS) PHOTOMOD (Russia) can be used as a local DPS with a distributed network environment [20]. DPS is the main tool for creating high-quality images for AGIS CENSs.

Also, let us mention the Talka technology and the digital photogrammetric station developed at Trapeznikov Institute of Control Sciences (Russian Academy of Sciences) [21]. At present, it does not support cloud computing: the technology is conventional, being implemented on the institute's server. Most of the world's leading websites are currently inaccessible for Russian software products. Cloud services are also unavailable, and there are no Russian cloud services in GIS technologies.

Over the past 50 years, computing has been based on five major platforms: mainframes (powerful computing systems), minicomputers, personal computers, servers, and mobile devices. Now a sixth type, the cloud platform, is emerging. Therefore, GIS developers consider cloud services and parallel computing an important line for developing their own platforms, including AGIS CENSs.

When choosing an appropriate platform for multiple information tasks, one should keep in mind that the gain in computational efficiency depends on the task algorithm and the number of consecutive calculations [8, 22]. For some tasks, increasing the number of processors in a computing system not necessarily leads to higher efficiency.

In this case, Amdahl's law concerns not the components of a single task but the set of tasks implemented in applied GISs to design search CENSs with given properties for efficient navigation.

The following algorithms and hypotheses are directly related to the final result of the navigation system [23] and are used when implementing CENS models:

– the types of Earth's physical fields (optical, thermal, radio-thermal, radar, etc.),
– possible image search and processing algorithms in onboard computers,
– weather conditions at the time of navigation,
– possible carrier's location errors,
– possible sizes of navigation sections,
– linear sampling intervals of the original map,
– navigation system altitude,
– pattern recognition algorithms,
– scene analysis algorithms,
– algorithms for determining navigation references,
– clustering algorithms,
– training algorithms for neural networks,
– possible values of the CENS decision function for accurate navigation and navigation with a given error,
– stress algorithms for the shooting system of CENSs with different means of stress exposure in EMR wavelength ranges with specified conditions (accuracy, the capabilities of such means, etc.).

The map of stress exposures on the shooting system of CENSs is built in accordance with the adopted algorithms and hypotheses for modeling CENSs. This map determines the conditions of accurate (specified) navigation and navigation with a given error. All the algorithms and hypotheses mentioned above are dynamic: they can be changed within the specified ranges.

Distributed GISs installed in several relevant-profile enterprises may provide some cloud services. These can be data prepared in advance, e.g., maps with navigation properties of a given area based on images in different EMR wavelength ranges. In particular, such maps may have reference points obtained using the SURF method. This method searches for reference points in images using the Hessian matrix and then creates their descriptors invariant to the image scale and rotation.

Thus, the general-purpose GIS platform should be extended and implemented in the form of AGIS CENSs. The same applied GIS serves to implement the mapping models of stress exposures on the shooting system using the reflection and absorption coefficients of terrain objects [19].

Finally, note serverless cloud environments in which a database adjusts itself to the load [24]. The result of event processing is independent of the server memory state. These elastic calculations may be of interest in modeling CENSs and need further research.

Serverless computing has emerged due to the transition from servers to virtualization and containerization: more resources are provided for tasks rather than platform and infrastructure maintenance [12]. Essentially, this approach is an implementation of multithreaded and parallel computing in cloud serverless environments. Serverless computing is not yet used in Russian enterprises; see the publications [16–18, 25]. Probably, serverless cloud environments need not be used in cloud GISs, particularly in applied GISs for modeling CENSs. All necessary information should be concentrated on the servers of distributed GISs of several profile organizations. This issue is debatable.

CONCLUSIONS

Cloud technologies, parallel computing, and their state-of-the-art and use in cloud GIS have been analyzed. The main conclusions are as follows.

• The dependence of Russian organizations on foreign technologies has been very strong over the last few years. Many foreign IT vendors have recently left Russia and have limited supply and support of licenses, including data storage systems, servers, and information protection systems. There is an obvious need for a comprehensive approach to building domestic infrastructure, from architecture to data backup and protection.

• When introducing cloud technologies in a cloud AGIS CENSs, it is necessary to find novel architecture and software solutions and alternatives among Russian suppliers and developers. At the same time, it is necessary to solve the problem of product compatibility.

• In cloud environments, the execution of tasks that were performed autonomously in GISs should be planned using parallel computing in AGIS CENSs with specified properties considering the entire EMR wavelength range in transparency windows, including disturbances in the form of stress exposures on the sensors of navigation systems to ensure effective navigation under interference.

• The peculiarities of the parallel structure of CENS search algorithms are revealed. When implementing these algorithms in parallel computing systems, proper consideration of the peculiarities allows utilizing their advantages to the highest degree. A computing system for modeling CENSs should provide parallelism at the level of command and data flows when implementing the models of search CENSs using cloud technologies and parallel computing.

• The components of AGIS CENSs with analogs in related problem domains have been identified. Under appropriate access to such components and data, the implementation costs of AGIS CENSs can be minimized. Access can be organized in a network of distributed GISs with the required level of information protection.

• It is possible to implement AGIS CENSs using classical technologies based on the existing image analysis and processing systems, cloud GISs, and general-purpose cloud technologies. At present, classical image processing technologies remain the most attractive. In this case, source materials must be used in strict compliance with the corresponding regulations; the materials used and created must be protected. General-purpose cloud technologies will still occupy an insignificant place in the tasks of modeling CENSs: the creation of new applications will require time and funding.

REFERENCES

1. Alchinov, A.I. and Gorokhovsky, I.N., A Conceptual Applied Geographic Information System for Modeling Search Autonomous Correlation-Extreme Navigation Systems, *Control Sciences*, 2022, no. 1, pp. 43–54.
2. Alchinov, A.I. and Gorokhovsky, I.N., Analysis of Stress Exposures on Autonomous Navigation Conditions in Search Correlation-Extreme Navigation Systems, *Control Sciences*, 2022, no. 6, pp. 35–48.
3. Marinescu, D.C., *Cloud Computing. Theory and Practice*, 2nd ed., Waltham: Morgan Kaufmann Publishers, 2018.
4. Voevodin, V.V. and Voevodin, Vl.V., *Parallel'nye vychisleniya* (Parallel Computing), St. Petersburg: BKhV – Peterburg, 2002. (In Russian.)
5. Ezhova, N.A. and Sokolinsky, L.B., Survey of Parallel Computation Models, *Bulletin of the South Ural State University. Series: Computational Mathematics and Software Engineering*, 2019, vol. 8, no. 3. pp. 58–91. DOI: 10.14529/cmse190304. (In Russian.)
6. Cloud Computing. URL: https://www.tadviser.ru/index.php/Cloud_Computing (Accessed March 29, 2023.)
7. Zhang, G., Zhu, A., and Huang, Q., A GPU-Accelerated Adaptive Kernel Density Estimation Approach for Efficient Point Pattern Analysis on Spatial Big Data, *International Journal of Geographical Information Science*, 2017, vol. 31, no. 10, pp. 2068–2097.
8. Zhuravlev, Yu.I. and Gurevich, I.B., Pattern and Image Recognition, in *Raspoznavanie, klassifikatsiya, prognoz. Matematicheskie metody i ikh primenenie* (Recognition, Classification, Forecast. Mathematical Methods and Their Application), Zhuravlev, Yu.I., Ed., Moscow: Nauka, 1989, vol. 2, pp. 5–72. (In Russian.)
9. Nupponen, J. and Taibi, D., Serverless: What It Is, What to Do and What not to Do, *Proc. Int. Conf. Softw. Archit. (ICSA 2020)*, Salvador, 2020, pp. 49–50. DOI: 10.1109/ICSA-C50368.2020.00016.
10. Serrano, N., Gallardo, G., and Hernantes, J., Infrastructure as a Service and Cloud Technologies, *IEEE Software*, 2015, vol. 32, no. 2, pp. 30–36.

11. Chernyak, L., Integration is Cloud Foundation, *Open Systems. DBMS*, 2011, no. 7. URL: https://www.osp.ru/os/2011/07/13010473/ (In Russian.)
12. Yang, C., Huang, Q., Li, Z., et al., Big Data and Cloud Computing: Innovation Opportunities and Challenges, *International Journal of Digital Earth*, 2016, vol. 10, no. 1, pp. 1–41. DOI: 10.1080/17538947.2016.1239771.
13. Ye, X., Huang, Q., and Li, W., Integrating Big Social Data, Computing and Modeling for Spatial Social Science, *Cartography and Geographic Information Science*, 2016, vol. 43, no. 5, pp. 377–378.
14. Huang, Q., Li, J., and Zhang, T., Domain Application of High Performance Computing in Earth Science: An Example of Dust Storm Modeling and Visualization, in *High Performance Computing for Geospatial Applications*, Cham: Springer, 2020, pp. 249–268.
15. Li, Z., Tang, W., Huang, Q., et al., Introduction to Big Data Computing for Geospatial Applications, *International Journal of Geo-Information*, 2020, vol. 9, no. 8, art. no. 487.
16. Google Cloud Tensor Processing Unit. URL: https://tadviser.com/index.php/Product:Google_Cloud_Tensor_Processing_Unit (Accessed October 10, 2022.)
17. Operation in GIS Panorama Using GeoCloud. URL: https://gisinfo.ru/techno/geocloud.htm (Accessed October 3, 2022.)
18. BIM in Cloud. A Survey of Key Information Modeling Solutions. URL: https://www.cloud4y.ru/blog/cloud-bim-solutions-overview/ (Accessed May 20, 2023.)
19. Lubnin, D.S., Modernization of the Spatial Data Infrastructure Using Cloud Technologies, *Izvestiya Vuzov. Geodesy and Aerophotosurveying*, 2018, vol. 62, no. 5, pp. 590–598. DOI 10.30533/0536-101X-2018-62-5-590-598. (In Russian.)
20. Sechin, A.Yu. and Adrov, V.N., Photogrammetry and Cloud Technologies, *Geoprofi*, 2017, no. 2, pp. 24–27. (In Russian.)
21. Alchinov, A.I., Beklemishev, N.D., and Kekelidze, V.B., *Metody tsifrovoi fotogrammetrii. Tekhnologiya "Talka"* (Methods of Digital Photogrammetry. The Talka Technology), Moscow: State University of Press, 2007. (In Russian.)
22. Cheremisinov, D.I., Amdahl's Law and Bounds on Speedup, *Proceedings of the 6th International Scientific and Practical Conference "BIG DATA and Advanced Analytics,"* Minsk, 2020, pp. 294–300. (In Russian.)
23. Yu, M., Bambacus, M., Cervone, G., et al., Spatiotemporal Event Detection: A Review, *International Journal of Digital Earth*, 2020, vol. 13, no. 1, pp. 1–27.
24. Lloyd, W., Ramesh, S., Chinthalapati, S., et al., Serverless Computing: An Investigation of Factors Influencing Microservice Performance, *Proc. Int. Conf. Cloud Eng. (IC2E)*, Orlando, 2018, pp. 159–169. DOI: 10.1109/IC2E.2018.00039.
25. Fomichev, A. and Bondar', O., A Serverless Alternative to Traditional Databases, *Open systems. DBMS*, 2021, no. 1, pp. 20–23. DOI: 10.51793/OS.2021.86.20.003. (In Russian.)

This paper was recommended for publication by V.B. Pavlov, a member of the Editorial Board.

Received April 24, 2023,
and revised July 5, 2023.
Accepted July 18, 2023.

Author information

Alchinov, Alexander Ivanovich. Dr. Sci. (Eng.), Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
✉ alchinov46@mail.ru
ORCID iD: https://orcid.org/0000-0002-6553-376X

Gorokhovsky, Igor Nikolaevich. Dr. Sci. (Eng.), Research Center of Topographic and Navigational Support, Central Research Institute No. 27, Moscow, Russia
✉ gin_box@mail.ru

Cite this paper

Alchinov, A.I., Gorokhovsky, I.N., Expanding the Functionality of an Applied Geographic Information System for Modeling Search Correlation-Extreme Navigation Systems. *Control Sciences* **5,** 65–74 (2023). http://doi.org/10.25728/cs.2023.5.6

Translated into English by *Alexander Yu. Mazurov*,
Cand. Sci. (Phys.–Math.),
Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
✉ alexander.mazurov08@gmail.com

DOI: http://doi.org/10.25728/cs.2023.5.7

COMPONENT MONITORING TO MANAGE THE REDUNDANCY OF AN ONBOARD EQUIPMENT COMPLEX

V.N. Bukov[1], A.M. Bronnikov[2], A.V. Vorob'ev[1], A.S. Popov[3], and V.A. Shurman[4]

[1] Institute of Aircraft Equipment, Zhukovsky, Russia
[2] Bauman Moscow State Technical University, Moscow, Russia
[3] Zhukovsky–Gagarin Air Force Academy, Voronezh, Russia
[4] Ramenskoe Instrument-Making Design Bureau, Zhukovsky Branch, Zhukovsky, Russia

✉ v_bukov@mail.ru, ✉ bronnikov_a_m@mail.ru, ✉ vorobiev@niiao.ru, ✉ saga30@yandex.ru, ✉ vshurman@rpkb.ru

Abstract. This paper considers the technical condition monitoring problem for the components of an onboard equipment complex to perform its real-time reconfiguration. The idea is to use at least three levels of monitoring systems: the nearest perspective, only traditional built-in control (BiC) to detect faults; the next level, BiC together with auxiliary means to increase the fidelity of technical diagnosis, including mutual cross-pair monitoring; the distant perspective, logical processing algorithms for system observations as a whole based on the normalized rules of failure mode and effects analysis (FMEA) of aircraft equipment. Mathematically, the pair monitoring of component conditions consists in forming the so-called preference matrices; their values and special tables are used to determine the condition of diagnosed objects with high reliability and, moreover, to evaluate possible errors of diagnostic tools. For third-level methods, an action sequence is proposed as follows: the reverse and direct logical models reproducing the dependencies of faulty states based on FMEA results are alternatively initiated. An updated methodology for handling triplex logical models is proposed. The main advantages of logical models—significant simplicity and universality—ensure their effectiveness in a wide range of dynamic systems of varying complexity. A methodological example illustrates the application of logical triplex models.

Keywords: onboard equipment complex, technical monitoring, logical pair monitoring, logical triplex models, analysis of functional failures, redundancy management.

INTRODUCTION

The creation of redundant reconfigurable onboard equipment complexes (OECs) for mobile objects is a non-alternative way to achieve the maximum possible reliability of these complexes under the limited reliability of the components used and a wide range of external factors. According to the paper [1], the concept of Active Fault-Tolerant Control Systems (AFTCS) implies the joint operation of at least three subsystems as follows. The first subsystem is the reconfigurable or adaptable part of the equipment (Reconfigurable Control System, RCS; in our case, the equipment of the complex); the second subsystem detects (monitors) and diagnoses faults of this complex (Fault Detection and Diagnosis, FDD); the third subsystem implements the so-called Reconfiguration Mechanism (RM).

Reconfigurability means that an OEC has the properties to purposefully change its parametric and structural characteristics in real time.

By definition, redundancy management is assigned to the second and third subsystems mentioned above.

Historically, research works on monitoring, diagnosis, and reconfiguration of technical systems have been isolated from each other. On the one hand, the known solutions in the field of monitoring and diagnosis [2–9] are not related to the subsequent use of their results in real time. On the other hand, the reconfiguration approaches proposed in [10–14] proceed from monitoring results as a given. This situation has obvious disadvantages since the following questions remain open: What is the actual necessity of real-time diagnosis? What are the mutual requirements of diagnosis and reconfiguration? How can their interaction be systematically analyzed?

Nevertheless, separate solutions of monitoring, diagnosis, and reconfiguration problems have been prevailing in the scientific literature and practice so far.

This paper is devoted to the technical condition (operability) monitoring problem for the components of a reconfigurable complex within the redundancy management approach based on the supervisory configuration control method [15]. Here, we consider a broader problem statement: the readiness monitoring of an OEC, which covers (along with operability) the completion of all real-time preparations of OEC components for the intended use.

1. THE FUNCTIONS AND GENERATIONS OF MONITORING MEANS

In the emerging applied theory of redundancy management [15], monitoring means have to determine, for each available (hardware or software) component, its availability index (AI) and functional efficiency (FEI) for use in periodic arbitration of configurations.

By assumption, monitoring is performed in three main steps as follows:

- data acquisition from components (either by sending special requests or intercepting data translation implemented by components independently);
- data processing, including initial mathematical treatment, if they come from several sources and belong to one component, and preparation of the results (generation of AI and FEI) for transmission;
- transfer of the results (AI and FEI) to the redundancy management level by supervisor requests or by translation over a shared local area network.

As the theory and applications evolve [15], different monitoring methods in terms of principles and algorithms can be used; they form three main levels summarized in Table 1.

The first monitoring level is basic and involves only the existing built-in control (BiC) or new means created by component developers. The effect is to ensure equipment control in accordance with industry regulations determining the depth and quality of control procedures [16, 17].

Along with the existing BiC, the second level involves effective algorithmic solutions with mutual control of redundant components having BiC, e.g., logical pair monitoring (LPM) procedures [18]. The corresponding implementation is possible in onboard automated control systems (OACSs) and onboard maintenance systems (OMSs) [17].

In the case of heterogeneous[1] objects and their BiCs, it is possible to assess the operability of the functional part of the components and BiC with maximum reliability.

The third level involves more complex (most importantly, independent of BiC) algorithms and monitoring strategies based on the analysis of processes in the “object + OEC” system using various concepts and models, including state forecasting. In particular, the matter concerns logical models based on the so-called directed triplex graphs [19].

Table 1

Development of monitoring methods for redundant OECs

No.	Level	Means	Toolkit	Effect
1	Basic	Traditional BiC	Autonomous monitoring	Ensuring equipment control in accordance with industry regulations
2	Next-generation	Integration with OACSs, use of LPM	Integration into the architecture of existing (currently developed) OACSs (OMSs)	More comprehensive and complementary equipment control based on various engineering and algorithmic solutions
3	Next-generation	Independent monitoring algorithms	Application of more sophisticated monitoring algorithms and strategies (logical models, state forecasting)	The qualitatively new level and high fidelity of control, multiple fault detection and diagnosis

[1] Nodes of the same purpose are created by different developers and (or) are based on different engineering solutions.

2. MONITORING BASED ON BUILT-IN CONTROL

BiC is a set of hardware or software components introduced into systems, their parts, or functional assemblies (FAs). As a rule, they do not participate in the work of functional modules (FMs) of the system or its FAs on purpose but collect and summarize various data that objectively reflect the operability of these modules in the developer's opinion.

There are two significantly different organizational approaches to the operation of BiC:

– test control of equipment operability, which requires a temporary "withdrawal" of the controlled object from its intended-purpose operation;

– functional control, which is performed during the intended-purpose operation of the controlled object.

Functional control is generally implemented based on two main principles as follows:

• *Use of different voting schemes.* Here, a common solution is the so-called quorum elements (QEs), which identify faulty modules by the processing of voting results of several connected FMs. The operability of an FM is judged by a significant deviation of its output from those of same-type modules (the largest deviation or that exceeding a given threshold) [20].

The main features of the quorum-based method include:

– the assumption that the technical state of an FM remains unchanged within each cycle;

– the assumption that a QE is operable (never fails);

– applicability to three or more FMs (in the case of two FMs, a pair of FMs becomes the controlled object, not each FM separately);

– the assumption that within the voting rules (equal, weighted, with discriminations, etc.), the operable FMs within each cycle dominate over the faulty ones and the latter can be disconnected;

– a common data flow for all FMs.

A peculiar form of voting is widely implemented in the so-called self-checking systems [9]: a set of same-type modules subjected to identical input actions is divided into pairs, and the outputs within each pair are compared with each other. A pair with matching outputs is considered to be operable; otherwise, both modules of the pair are considered to be inoperable.

• *Use of fidelity rules.* Depending on particular conditions and solutions, such rules can be as follows: comparing with reference models, detecting violations of given time and (or) parametric intervals (control by parameter tolerance [20]), checking logical and other relations, calculating different-order invariants, etc.

The main features of the method of fidelity rules include the following:

– Within each cycle, the operability of an FM does not change.

– By assumption, an element implementing fidelity rules is operable. (If there is a reference model, it is operable.)

– This method is applicable to any number of FMs.

– By assumption, the input and output data contain sufficient information.

– Each FM has a separate data flow.

The technical condition of computing units in the OEC central computing system is monitored by combining the methods described above.

In accordance with the ARINC 653 standard, a health monitor is a system function responsible for monitoring and reporting errors in the operation of hardware means, application software, and the operating system. The information about the technical condition of the computer during normal operation is finally collected by operating system kernel mechanisms and (or) a special section of system software.

The resulting information of the health monitor is transmitted to the OMS and communication channels with ground facilities or is processed by the operating system. This information is generated from the following input data:

• the results of built-in control tests, which check equipment operability in the background mode during specially allocated time intervals;

• the output data of event handlers; an event is a "special case" detected by hardware means when executing functional applications, usually a programming error detected or a protocol violation during data reception in the input channels of an external interface;

• the information of functional applications about errors and incorrect input or output data.

The recent direction [21] stands somewhat apart. It can be called FM monitoring based on operational data. By assumption, a special element (chip) is structurally and functionally connected directly to an FM to gather and accumulate data on the conditions of its use and storage. Such a chip stores different parameters (FM data) and sends them to the monitoring module, in particular:

– passport information,

– test results at different stages of the life cycle,

– statistics of operation indicators and characteristics (estimates of the achieved accuracy, remaining life, energy indicators, etc.),

– statistics of external impacts during intended use, storage, and routine maintenance.

The monitoring module is responsible for analyzing the incoming data and judging about FM operability based on the analysis results.

Thus, we summarize the common features (limitations) of BiC with different degrees of occurrence:

– weak[2] assumptions about the unchanged operability of the controlled devices within the monitoring cycle;

– strong[3] assumptions about the operability of control systems or their major devices;

– the requirement on a minimum admissible or large number of FMs (in the case of quorum or majority control);

– the requirement that operable FMs dominate over inoperable FMs;

– the fast disconnection of faulty FMs;

– the requirement on sufficient informativeness for all processes in FMs.

The main advantage of using BiC (in the current form) to monitor the components of a redundant OEC is the well-established technologies of their creation and application in practice.

Analytical monitoring and diagnosis methods [5–7] are being intensively developed. They further refine the concept of fidelity rules and are based on theoretical patterns and peculiarities in the operation of dynamic systems.

3. USE OF LOGICAL PAIR MONITORING

A common drawback of using BiC is the forced trust in these diagnosis means, i.e., the a priori assumption of their infallibility [22–24]. According to the studies [15], without considering the inevitable limited capabilities of control (monitoring) means, factual fault tolerance can be significantly inferior to expectations.

One remedy is logical pair monitoring (LPM) procedures [18]: both autonomous monitoring and mutual cross-monitoring are performed for two FMs of the same functional purpose. By assumption, all structurally isolated functional assemblies "FM + BiC," comparable in purpose and operation principles, are designed so that the BiC of each assembly can access the FM of any other assembly[4] (Fig. 1).

This figure has the following notations: τ is the current time (monitoring cycle number); v_τ is input data; y_τ is output data; p_τ is controlled parameters (possibly, they include v_τ and y_τ); s_τ^{i-j} is the operability assessment of the i th FM (FMi) generated by the j th BiC (BiCj). Binary operability assessments (1—"operable" and 0—"inoperable") form an indicator matrix (IM).

Monitoring is performed under several assumptions:

- Same-type data flows through different FAs are not connected with each other (the functional autonomy of FAs).
- Each functional node "FM + BiC" is implemented on a technological base and supported by infrastructural means independently of the base and means of other FAs (the technical heterogeneity of FAs).
- FMs may be independently operable or inoperable (the independence of FM operability).
- Only one element in a pair of BiC can have a simple error, i.e., a false assessment of "operable" or "inoperable" (the error-free operation of at least one BiC).
- The monitoring process is divided into cycles within which FM operability and the errors of BiC are unchanged (the stationary operability of FAs).

Under these assumptions, the full group of different IM values makes up 13 matrices uniquely related to the operable or inoperable state of both FAs; for details, see [18]. In accordance with the LPM indicator rule [18] (Table 2 below), each IM value unambiguously determines the technical condition of both each FM and each BiC. The only exception is the IM value

$$S_\tau^{\text{ind}} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$

which describes the operability of both FMs or the inoperability of one BiC in the form of a false assessment of "operable." But this ambiguity does not concern FM operability; in the part of BiC, it can be considered by design solutions.

Also, the publication [25] presented a more complicated version of LPM with the possible presence of the so-called gray zone. This zone appears when some part of BiC participating in monitoring cannot be separated from the FM in terms of data passage. In this case, to implement LPM, BiC has to be dissected along the boundary between the grey zone and the analytical segment. As a result, the indicator rule is modified and, generally, the efficiency of monitoring is reduced: the operability of the grey zone becomes indistinguishable from that of the FM, and error identification refers to the analytical segment only.

---

[2] This assumption is not crucial in practice.

[3] This assumption significantly narrows the applicability of the approach.

[4] The complete adoption of such an idea may cause significant difficulties. As a compromise solution, limited access may be implemented at the developer's discretion.

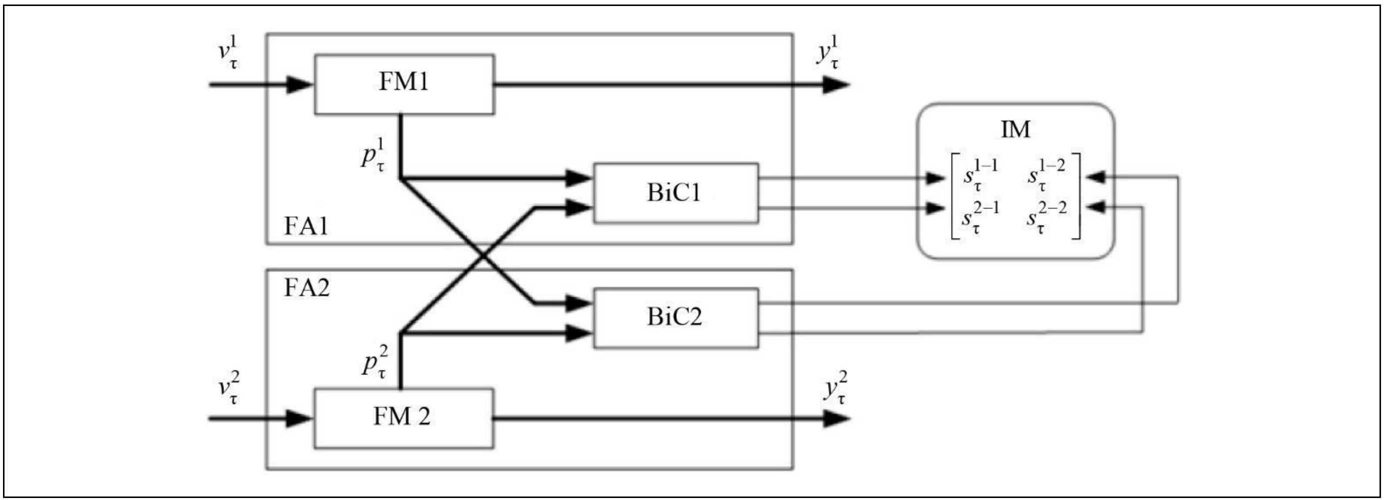

Fig. 1. The schematic diagram of functional assemblies for LPM.

Table 2

IM values obtained by LPM

Inoperable FM	Errors in BiC				
	Errors in BiC1		Errors in BiC2		No errors
	False “1”	False “0”	False “1”	False “0”	
Inoperable FM1	$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$
Inoperable FM2	$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$
Both operable	–	$\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$	–	$\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

4. DIRECT AND REVERSE LOGICAL MODELS

The publications [20, 26] introduced a logical models-based approach to operability control of technical systems. When applied to the monitoring of components in a complex with managed redundancy, this approach is as follows.

According to [27], the OEC designer performs failure mode and effects analysis (FMEA) of aircraft equipment. Such failures include completely terminated operation, loss of capability to meet the requirements, intermittent operation, unnecessary operation, etc. The assessment results are a list of typical equipment failures causing functional faults, including the description of their relationships and consequences. In most cases, failure consequences are divided into local (i.e., characteristic of the component itself), those of the next higher level, and those at the highest level (the entire system, e.g., an aircraft). It is necessary to identify failure consequences at the highest level for comparing the criticality of failures of all components included in the OEC. Usually, FMEA results are presented in the form of tables of possible failures and their consequences.

FMEA may be not comprehensive, i.e., may be performed partially considering the criticality of failures of different parts of the systems.

This approach implies passing from descriptive (qualitative) FMEA results to the construction of two types of formalized logical failure propagation models for the object of diagnosis with triplex variables: 0—“no failure or its impact,” 1—“failure or its impact exists,” and &—“indefinite state.”

The original methodology of building and using triplex models was described in [20, 26]. It has a low level of formalization (a system of decision rules), which creates difficulties in practice. Below we propose a deeper approach largely devoid of this drawback.

The idea is to model failure impact propagation (from causes to manifestations) in an OEC using a logical network containing generalized elements with the following logical operators: ORi (an analog of disjunction) or ANDi (an analog of conjunction) at the input and ORo or ANDo at the output; see Fig. 2. In this case, the element's state x_{id} is given by the values of triplex variables at its inputs, x_{in}^{j}, in accordance with causal relation formulas: $x_{\text{in}}^{1} + x_{\text{in}}^{2} \to x_{\text{id}}$ for the ORi operator or $x_{\text{in}}^{1} \times x_{\text{in}}^{2} \to x_{\text{id}}$ for the ANDi operator. It also determines the value of such variables at the outputs: $x_{\text{id}} \to x_{\text{out}}^{1} + x_{\text{out}}^{2}$ for the ORo operator or $x_{\text{id}} \to x_{\text{out}}^{1} \times x_{\text{out}}^{2}$ for the ANDo operator.

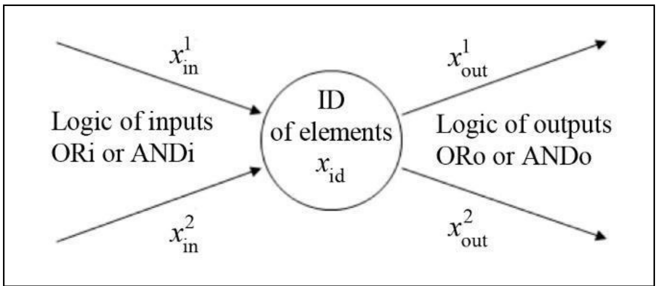

Fig. 2. An element of the logical network (failure impact propagation models for OEC).

The ORo operator must be provided with a description of output switching conditions (by an external impact, by definite characteristics of the logical network state, etc.).

Note an important aspect: when constructing the logical network by the artificial division of the models of real devices, each network element must be assigned at most one input and at most one output operator.

Note that the arithmetic of such triplex variables is not conventional, and each causal transition between the elements of the proposed logical network (the direct logical model) is performed according to the rules summarized in Table 3. The additional symbol ∇, also called an operator, denotes the absence of an alternative.

For example, the formulas in cells 1-2 (row 1, column 2) and 1-4 of Table 3 should be interpreted as follows: the presence of at least one signal $x_{\text{in}}^{j} = 1$ (failure impact) at the inputs of a logical network element with the ORi logic brings this element to the state $x_{\text{id}} = 1$ (it is prone to failure impact). This corresponds to failure state evolution in non-redundant functional devices of the OEC. The cells of row 2, on the other hand, are associated with the ANDi logic, characteristic of redundant devices.

The cells highlighted in yellow in Table 3 describe the propagation of the indefiniteness & over the triplex logical network. The remaining cells reflect the unambiguous development of the situation: the propagation (1) or non-propagation (0) of failure impact.

The process of analyzing system failures is associated with the reverse logic that determines transitions from failure manifestations to their causes. The corresponding transitions (the reverse logical model, the left "effect-to-cause" arrows) $x_{\text{in}}^{1} + x_{\text{in}}^{2} \leftarrow x_{\text{id}}$, $x_{\text{in}}^{1} \times x_{\text{in}}^{2} \leftarrow x_{\text{id}}$, and $x_{\text{id}} \leftarrow x_{\text{out}}^{1} \times x_{\text{out}}^{2}$ are presented in Table 4. For example, the formula in cell 2-1 of this table is interpreted as follows: the state $x_{\text{id}} = 1$ of an element with the ANDi operator was a consequence of the simultaneous presence of 1 at its inputs.

Reverse logic is used to judge about the input given a known output. For operators at the input, it is required to determine possible combinations at the input of a logical network element by its state; for operators at the output, it is required to determine the state of a logical network element by the combination at the output.

Table 3

Direct logic arithmetic

Operators and row numbers		Column numbers and formulas								
		1	2	3	4	5	6	7	8	9
ORi	1	$1+1\to 1$	$1+0\to 1$	$1+\&\to 1$	$0+1\to 1$	$0+0\to 0$	$0+\&\to \&$	$\&+1\to 1$	$\&+0\to \&$	$\&+\&\to \&$
ANDi	2	$1\times 1\to 1$	$1\times 0\to 0$	$1\times \&\to \&$	$0\times 1\to 0$	$0\times 0\to 0$	$0\times \&\to 0$	$\&\times 1\to \&$	$\&\times 0\to 0$	$\&\times \&\to \&$
ORo	3	-	$1\to 1+0$	-	$1\to 0+1$	$0\to 0+0$	$\&\to 0+\&$	-	$\&\to \&+0$	-
ANDo	4	$1\to 1\times 1$			$0\to 0\times 0$			$\&\to \&\times \&$		
∇	5	$1\to 1$			$0\to 0$			$\&\to \&$		

Table 4

Reverse logic arithmetic

Operators and row numbers		Column numbers and formulas		
		1	2	3
rORi	1	0 + 0 ← 0	(1 + 1 ← 1 or 1 + 0 ← 1 or 0 + 1 ← 1 or)* 1 + & ← 1 or & + 1 ← 1	0 + & ← & or & + 0 ← & or & + & ← &
rANDi	2	1 × 1 ← 1	(0 × 0 ← 0 or 1 × 0 ← 0 or 0 × 1 ← 0 or)* 0 × & ← 0 or & × 0 ← 0	1 × & ← & or & × 1 ← & or & × & ← &
rORo	3	0 ← 0 + 0	1 ← 1 + 1 or 1 ← 1 + 0 or 1 ← 0 + 1 or 1 ← 1 + & or 1 ← & + 1	& ← 0 + & or & ← & + 0 or & ← & + &
rANDo	4	1 ← 1 × 1	0 ← 1 × 0 or 0 ← 0 × 1 or 0 ← 0 × 0 or 0 ← & × 0 or 0 ← 0 × &	& ← 1 × & or & ← & × 1 or & ← & × &
∇	5	1 ← 1	0 ← 0	& ← &

* If "indefinite" is conceptually identified with "any," then the formulas in brackets should be ignored.

The formulas in Table 4 can be justified using the following explanations for row 4:

a) If a combination of 1 and 1 is detected at the outputs of a logical network element (both are prone to failure impact), then due to the logic of the AND operator, this element is prone to failure impact.

b) If a combination of 1 and 0 is detected at the outputs of a logical network element (one output is prone to failure impact, whereas the other is not), then this element is resistive to failure impact, and the failure has occurred in the chain of elements following the output with a value of 1.

c) If a combination of 0 and 1 is detected, then the result is the same as in item *b*).

d) If a combination of 0 and 0 is detected, then the element is resistive to failure impact.

e) If a combination of & and 0 is detected, then the element is resistive to failure impact for any value of & (see item *a*) or *c*)).

f) If a combination of 0 and & is detected, then the result is the same as in item *e*).

g) If a combination of & and & is detected, then the element has the indefinite state.

h) If a combination of & and 1 is detected, then the element has the indefinite state since the element is resistive to failure impact for & = 0 (by item *b*)) but is prone to failure impact for & = 1 (by item *f*)).

i) If a combination of 1 and & is detected, then the result is the same as in item *h*).

In Table 4, the cells generating ambiguity are highlighted in yellow; it is therefore required to analyze in parallel each of the possible variants. For example, according to the formulas in cell 1-2 (analysis of rORi, i.e., the ORi operator in the reverse direction), the state $x_{id} = 1$ may be a consequence of signal indefiniteness at any input, even though the other input may be resistive to failure impact. Bold boxes indicate cells with different combinations of output signals corresponding to the same element state. For example, according to cell 3-2, an element with the ORo operator is prone to failure impact if any of its outputs has this property.

The formulas of Tables 3 and 4 can be obviously extended to the cases of three or more inputs and outputs.

5. FORMULAS OF TRIPLEX MODELS

The direct analysis of the logical network is to model cycle-to-cycle failure impact propagation among elements. The corresponding relations of state dynamics and output have the form

$$X_{k+1} = \mathrm{DM} \overset{\rightarrow}{\Diamond} X_k + X_{\mathrm{init}},\ Y_k = \mathrm{EM} \times X_k,\quad k = 1, 2, \ldots, \tag{1}$$

with the following notations: X_k is the *n*-dimensional vector of OEC component failures at calculation cycle k with values 1, 0, or &, assigned to each logical network element; X_{init} is the initial state of the vector X_k; DM is the direct dependency matrix, filled with unities and empty elements[5] ⊙ according to FMEA results; EM is the exit matrix, which highlights the elements of the OEC model with directly observed failures (is filled with zeros and unities); Y_k is the *m*-

[5] Not a zero value (no failure), but an indication of being eliminated from consideration.

dimensional output vector of directly observed failures (or their absence). Here, the signs $\times$ and $+$ indicate the extended conjunction and disjunction, respectively (Table 3); the sign $\vec{\Diamond}$ means their simultaneous use according to the special methodology described in Section 6. The first formula in (1) is not algebraic in the conventional sense.

To explain the expression (1), we consider an illustrative example with five elements. The state of one element is directly observed. Figure 3 shows the graph of the corresponding logical network.

The graph of this example is described by the following relations (1):

$$\underbrace{\begin{bmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{bmatrix}_{k+1}}_{X_{k+1}} = \underbrace{\begin{matrix}\nabla\\ \text{ANDi}\\ \text{ORi}\\ \nabla\\ \nabla\end{matrix}\begin{bmatrix}\odot & \odot & \odot & 1 & \odot\\ 1 & \odot & \odot & 1 & \odot\\ 1 & 1 & \odot & \odot & \odot\\ \odot & \odot & 1 & \odot & \odot\\ 1 & \odot & \odot & \odot & \odot\end{bmatrix}_{\begin{matrix}\text{ANDo} & \nabla & \nabla & \text{ORo} & \nabla\end{matrix}}}_{\text{DM}} \vec{\Diamond}$$

$$\vec{\Diamond}\underbrace{\begin{bmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{bmatrix}_k}_{X_k} + \underbrace{\begin{bmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{bmatrix}_{\text{HO}}}_{X_{\text{HO}}}, \qquad (2)$$

$$\underbrace{\begin{bmatrix}y\end{bmatrix}_k}_{Y_k} = \underbrace{\begin{bmatrix}1 & 0 & 0 & 0 & 0\end{bmatrix}}_{\text{EM}} \times \underbrace{\begin{bmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{bmatrix}_k}_{X_k},$$

in DM, the corresponding input operators are given to the left of the rows, and the corresponding output operators are given under the columns; in addition, x_i denotes the states of logical network elements with values 1, 0, or &.

The cyclic use of formulas (1) (in the illustrative example, formulas (2)) for $k = 0, 1, 2, \ldots$ allows modeling the cycle-to-cycle impact propagation of the initial failure X_{init} over all logical network elements.

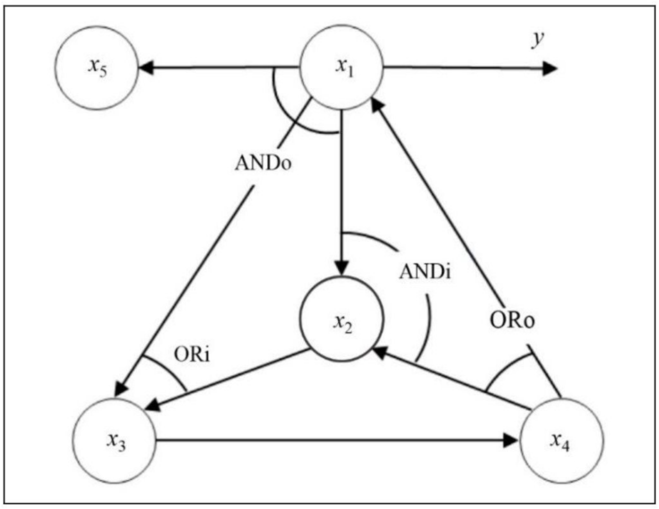

Fig. 3. An illustrative example of a logical network.

The reverse analysis is to find the root cause of failures by their manifestations. The corresponding iterative formulas for the initial and subsequent assessments have the general form

$$\hat{X}_0 = \text{EM}^{\text{T}} \times Y_k + \overline{\text{EM}}^{\text{R}} \times \mu(\&),$$

$$\hat{X}_{\tau+1} = \text{rDM}\,\overleftarrow{\Diamond}\,\hat{X}_\tau + \text{EM}^{\text{T}} \times Y_k,\ \tau = 0, 1, 2, \ldots, \qquad (3)$$

with the following notations: $\hat{X}_\tau$ is the estimated component failure vector at cycle τ with the initial estimate $\hat{X}_0$; $\overline{\text{EM}}^{\text{R}}$ is the matrix right divisor of zero of maximum rank for EM [28]; $\mu(\&)$ is a matrix of compatible dimensions with arbitrary elements, denoted by &. Here, the transposition of the binary matrix EM with linearly independent rows replaces the canonizer (a more complex universal matrix structure [28]), and the sign $\overleftarrow{\Diamond}$ means the reverse analysis operations for failure impact propagation (see the formulas in Table 4), which are performed using the special methodology described in Section 6.

The reversed dependency matrix (rDM) is obtained from DM by applying transposition, replacing the operators ORi $(x^1_{\text{in}} + x^2_{\text{in}} \to x_{\text{id}})$, ANDi $(x^1_{\text{in}} \times x^2_{\text{in}} \to x_{\text{id}})$, ORo $(x_{\text{id}} \to x^1_{\text{out}} + x^2_{\text{out}})$, and ANDo $(x_{\text{id}} \to x^1_{\text{out}} \times x^2_{\text{out}})$ by the operators rORi $(x^1_{\text{in}} + x^2_{\text{in}} \leftarrow x_{\text{id}})$, rANDi $(x^1_{\text{in}} \times x^2_{\text{in}} \leftarrow x_{\text{id}})$, rORo $(x_{\text{id}} \leftarrow x^1_{\text{out}} + x^2_{\text{out}})$, and rANDo $(x_{\text{id}} \leftarrow x^1_{\text{out}} \times x^2_{\text{out}})$, respectively, and adding 1 to the diagonal position of each row corresponding to the null row[6] of the combined matrix $\begin{bmatrix} \mathrm{DM}^{\mathrm{T}} & \mathrm{EM}^{\mathrm{T}} \end{bmatrix}$.

Thus, the direct logical model (2) relates to the reverse logical model (3) of the form

$$\underbrace{\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \\ \hat{x}_4 \\ \hat{x}_5 \end{bmatrix}_0}_{\hat{X}_0} = \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{\mathrm{EM}^{\mathrm{T}}} \times \underbrace{y_k}_{Y_k} + \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\overline{\mathrm{EM}}^{\mathrm{R}}} \times \underbrace{\begin{bmatrix} \& \\ \& \\ \& \\ \& \end{bmatrix}}_{\mu(\&)}, \quad (4)$$

$$\underbrace{\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \\ \hat{x}_4 \\ \hat{x}_5 \end{bmatrix}_{\tau+1}}_{\hat{X}_{\tau+1}} = \quad (5)$$

$$= \underbrace{\begin{matrix} \mathrm{rANDo} \\ \nabla \\ \nabla \\ \mathrm{rORo} \\ \nabla \\ \\ \end{matrix} \begin{bmatrix} \odot & 1 & 1 & \odot & 1 \\ \odot & \odot & 1 & \odot & \odot \\ \odot & \odot & \odot & 1 & \odot \\ 1 & 1 & \odot & \odot & \odot \\ \odot & \odot & \odot & \odot & \mathbf{1} \end{bmatrix}_{\begin{matrix} \nabla & \mathrm{rANDi} & \mathrm{rORi} & \nabla & \nabla \end{matrix}}}_{\mathrm{rDM}} \overleftarrow{\Diamond} \underbrace{\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \\ \hat{x}_4 \\ \hat{x}_5 \end{bmatrix}_{\tau}}_{\hat{X}_{\tau}} + \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{\mathrm{EM}^{\mathrm{T}}} \times y_k,$$

where bold type sets off the diagonal element 1 added by the rDM formation rule.

The peculiarities[7] of handling the matrices DM and rDM are discussed in Section 6.

6. THE METHODOLOGY FOR HANDLING TRIPLEX MODELS

In the direct logical failure impact propagation model (1), the operation $\vec{\Diamond}$ remotely resembles the multiplication of a square matrix by a column matrix on the right. These operations are not identical due to binding different rows and columns of the matrix DM to different input and output logical operators of the logical network and the presence of empty elements (see model (2) in the illustrative example).

Let $\mathrm{DM}_{i.j}$ denote the binary element of the *i*th row and *j*th column of the matrix DM. In the direct logical model, the operation $\vec{\Diamond}$ implies the sequential composition[8] of the elements $x_{j.k}$ of the column matrix X_k with the elements $\mathrm{DM}_{i.j}$ for each i. The element $\mathrm{DM}_{i.j} = 1$ corresponds to using the triplex value of the variable $x_{j.k}$ whereas $\mathrm{DM}_{i.j} = \odot$ to ignoring the latter. An operator associated with a row of the matrix DM prescribes the type of formulas from Table 3, combining the elements of this row. An operator associated with a column of the matrix DM prescribes preliminary actions with the triplex variable as follows: the operators ∇ and ANDo prescribe no actions; the operator ORo prescribes introducing the distinction of the variables $x_{j.k}$ used in this column according to the switching conditions in the logical network. These compositions of each row are combined according to the operators indicated for the rows of DM.

For example, the first formula in (2) is equivalent to

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}_{k+1} = \begin{matrix} \nabla \\ \mathrm{ANDi} \\ \mathrm{ORi} \\ \nabla \\ \nabla \end{matrix} \begin{bmatrix} \odot & \odot & \odot & x_4(\mathrm{ORo}) & \odot \\ x_1 & \odot & \odot & x_4(\mathrm{ORo}) & \odot \\ x_1 & x_2 & \odot & \odot & \odot \\ \odot & \odot & x_3 & \odot & \odot \\ x_1 & \odot & \odot & \odot & \odot \end{bmatrix}_k + \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}_{\mathrm{init}} = \begin{bmatrix} x_{4.k}(\mathrm{ORo}) + x_{1.\mathrm{init}} \\ x_{1.k} \times x_{4.k}(\mathrm{ORo}) + x_{2.\mathrm{init}} \\ x_{1.k} + x_{2.k} + x_{3.\mathrm{init}} \\ x_{3.k} + x_{4.\mathrm{init}} \\ x_{1.k} + x_{5.\mathrm{init}} \end{bmatrix}. \quad (6)$$

[6] A null row contains empty and zero elements only.
[7] They express the novelty of the methodology presented.
[8] This term is used here in the universal sense for extended conjunctions and disjunctions.

Here, the notation $x_{4.k}(\text{ORo})$ refers to cycle k and is interpreted as follows: according to the switching rule at the output of element 4 (Fig. 3), the current value $x_{4.k}$ is considered in either the first $(x_{1,k+1})$ or second $(x_{2,k+1})$ row of (6) only. The alternative to the value $x_{4.k}$ is 0.

With $x_{1.k} = \&$, $x_{2.\text{init}} = 1$, and $x_{4.k} = 1$, the row for $x_{2.k+1}$ gives

$$\text{for } x_4 \xrightarrow{\text{ORo}} x_1,\ x_{2.k+1} = \underbrace{\underbrace{(\& \times \mathbf{0})}_{0\ (3/2\text{-}6)} + 1}_{1\ (3/1\text{-}4)} = 1;$$

$$\text{for } x_4 \xrightarrow{\text{ORo}} x_2,\ x_{2.k+1} = \underbrace{\underbrace{(\& \times \mathbf{1})}_{\&\ (3/2\text{-}7)} + 1}_{1\ (3/1\text{-}7)} = 1.$$

Hereinafter, the subscripts in brackets indicate the cell address: table/row-column.

In the reverse logical failure cause analysis model (3), the operation $\overleftarrow{\Diamond}$ is interpreted as follows. Sequential composition is performed for the elements $\hat{x}_{j.\tau}$ of the column matrix $\hat{X}_\tau$ with the elements $\text{rDM}_{i.j}$ for each i. The element $\text{rDM}_{i.j}$ = 1 corresponds to using the triplex value of the variable $\hat{x}_{j.\tau}$ whereas $\text{rDM}_{i.j}$ = $\odot$ to ignoring the latter (but according to other rules). An operator associated with a column of the matrix rDM prescribes the type of formulas from Table 4. An operator associated with a row of the matrix DM prescribes combining the elements of this row and preliminary actions with the triplex variable as follows: the operators ∇ and rANDo prescribe no actions; the operator rORo prescribes introducing the distinction of the variables $x_{j.\tau}$ used in this column according to the switching conditions in the logical network.

The actions with rDM are explained by the following generalized notation, valid for any element rDM_{ij}:

$$\left[\hat{x}_i\right]_{\tau+1} = {}_{\text{rOPo}}\left[\ldots 1_{ij} \ldots\right] \overleftarrow{\Diamond} \left[\hat{x}_j\right]_\tau + \ldots$$
$$\text{rOPi} \qquad (7)$$
$$= \left[\ldots {}_{\text{rOPo}} \quad (?_{\text{rOPi}}? \leftarrow \hat{x}_{j.\tau})_{\text{table 4}} \quad {}_{\text{rOPo}}\ldots\right] + \ldots,$$

where rOPi outside square brackets denotes the operators rORi, rANDi, and ∇ at the input of the element, and rOPo outside square brackets denotes the operators rORo, rANDo, and ∇ at the output of the element. Inside square brackets, these notations represent the corresponding operations: + (in the case of OR), × (in the case of AND), or no operations (in the case of ∇). Question marks indicate the values 1, 0, or & read from Table 4 for particular rOPi and $\hat{x}_j$.

For example, the expression (8) is equivalent to formula (5). Here, the notations $x_{1.\tau}(\text{rORo})$ and $x_{2.\tau}(\text{rORo})$ are interpreted as follows: according to the switching rule at the output of element 4 (Fig. 3), either the value $\left(\hat{x}_{1,\tau}\right)$ or the value $(\hat{x}_{2,\tau})$ is used in the line for $\hat{x}_{4.\tau+1}$. The alternative is the value 0.

The ambiguities in formula (8) are resolved by harmonizing the logical formulas or require additional investigation of the variants.

$$\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \\ \hat{x}_4 \\ \hat{x}_5 \end{bmatrix}_{\tau+1} = \begin{bmatrix} \odot & \times & ?\times ? \leftarrow \hat{x}_{2.\tau} & \times & ?+? \leftarrow \hat{x}_{3.\tau} & \times & \odot & \times & \hat{x}_{5.\tau} \\ \odot & & \odot & & ?+? \leftarrow \hat{x}_{3.\tau} & & \odot & & \odot \\ \odot & & \odot & & \odot & & \hat{x}_{4.\tau} & & \odot \\ \hat{x}_{1.\tau}(\text{rORo}) & + & ?\times ? \leftarrow \hat{x}_{2.\tau}(\text{rORo}) & + & \odot & + & \odot & + & \odot \\ \odot & & \odot & & \odot & & \odot & & \hat{x}_{5.\tau} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \times y_k. \qquad (8)$$

With $\hat{x}_{2.\tau} = 1$, $\hat{x}_{3.\tau} = 0$, and $y_k = 1$, the row for $\hat{x}_{1.\tau+1}$ gives

$$\hat{x}_{1.\tau+1} = \underbrace{\underbrace{\underbrace{1\times 1 \leftarrow 1}_{1\ (4/2-1)} \times \underbrace{0+0 \leftarrow 0}_{0\ (4/1-1)}}_{0\ (4/4-2)} + 1}_{1\ (4/3-2)} = 1,$$

which agrees with the direct analysis of the graph in Fig. 3.

7. THE GENERAL LOGIC OF USING TRIPLEX MODELS

The failures of OEC components during their operation are diagnosed as follows. The proposed approach is based on the assumption that any technical system (the object of diagnosis) contains three groups of parts.

The first group includes various OEC components (hardware and software) whose failures are significant (critical) and are covered by FMEA.

Communication channels can be either physical (wired or wireless communication between components) or virtual (routed digital communication).

The third group includes OEC components to identify (observe) the correctness or incorrectness of OEC operation directly. As a rule, the final effects of functional failures are manifested in such components.

The three groups of components are characterized in Table 5.

The proposed approach is to use the direct and reverse logical models alternately and repeatedly. It has the following features:

Table 5

The capabilities and features of logical models

<table>
<tr><th colspan="2" rowspan="2">Characterization</th><th colspan="3">The group of system parts</th></tr>
<tr><th>System components where failures may occur</th><th>Links between components (the propagation channels of failure impacts)</th><th>The locations of failure manifestations (the devices whose behavior can be observed to detect the occurrence of failures)</th></tr>
<tr><td colspan="2">Features regarding the occurrence and manifestation of failures</td><td>Failures can be in any of the components analyzed.</td><td>The links can be arbitrary within known structures.</td><td>The locations of failure manifestations, as well as the forms of these manifestations, are precisely known.</td></tr>
<tr><td rowspan="2">The capabilities of logical models</td><td>Direct logical model</td><td>As a rule, the locations of expected failures (the vector X_k) are unknown and are given approximately.</td><td>The links must be determined precisely.</td><td>The locations of failure manifestations are calculated, but they possibly differ from the really observed ones due to erroneous specification of failure locations. Therefore, the modeling process is repeatedly initiated with varying the expected failures. The accuracy criterion of failure specification is the coincidence of the calculated and measured output vectors Y_k.</td></tr>
<tr><td>Reverse logical model</td><td>Failure locations are determined from calculation results, but there is no confidence due to model ambiguities.</td><td>The links are determined by the logical reversion of the direct model.</td><td>The locations of failure manifestations are specified according to the observation results.</td></tr>
</table>

• Due to its exceptional simplicity, the logical failure impact propagation model in the form of a logical network yields computationally simple algorithms even for very complex OEC architectures.

• Model building is based on FMEA, a well-mastered methodology for the aircraft industry with acceptable depth and breadth of coverage for the operation conditions of OECs.

• The operability of components is described using triplex variables to reduce the number of variants under analysis when executing the algorithms.

• The models are alternated to proceed with the following steps.

Step 0. The initial failure vector estimate $\hat{X}_{\tau=0}$ is determined by formula (3) from the known output vector Y_0 (the vector of directly observed failures). The estimate $\hat{X}_{\tau=0}$ contains the components of the vector Y_0 in the form of 0 and 1; the other components are indefinite.

Step 1. The reverse logical model is used from the locations of failure manifestations $\hat{X}_{\tau=0}$ to expected failures to divide the components $\hat{X}_{\tau=1,\ 2,...}$ into definitely operable (0), definitely inoperable (1), and indefinite (&). The indefinite states either pass through the branches of the reverse logical model unchanged or change to definite states. The number τ of cycles implemented must be sufficient to reach a "stationary point," i.e., the invariable vector $\hat{X}_{\tau}$.

Step 2. The direct logical model is used from the expected failures $\hat{X}_{k=0} = \hat{X}_{\tau}$ to the corresponding estimates $\hat{Y}_{k=1,2,...} = \text{EM} \times \hat{X}_{k=1,2,...}$ of their manifestations to confirm the adequacy of these estimates or to refine the indefinite states. The number *k* of cycles implemented must be sufficient to reach a "stationary point," i.e., the invariable vector $\hat{X}_{k=1,2,...}$.

Steps 1 and 2 are alternated until stabilizing the estimate $\hat{X}_{k=1,2,...}$.

The problem of calculating the estimate $\hat{X}_{k=1,2,...}$ from the vector Y_0 has no analytical solution so far. Hence, it can be obtained by numerical iterative methods only. Various computational algorithms, rational and effective to a greater or lesser extent, can be applied here. One possible algorithm was presented in [19].

8. A METHODICAL EXAMPLE

In [29], the failures in helicopter altitude and speed parameters were detected using the algorithm [19]. The paper [30] considered the problem of detecting failures in a redundant electrohydraulic actuator. Due to the voluminous nature of applications and the limited scope of the presentation, we will demonstrate the approach of this paper on a simplified example.

Consider a partially redundant OEC fragment (Fig. 4). Here, drive 1 is controlled by computing units 1 and 2 (the drive becomes inoperable in the case of failures of both units); drives 2 and 3 are controlled by computing unit 2; power unit 1 feeds computing units 1 and 2; power unit 2 feeds drives 1 and 2; finally, power unit 3 feeds drive 3. For the sake of a compact problem statement, neither power supply buses nor digital and analog communication lines are considered. The directly observable data are the drive failures, which are assessed by the current positions of the rods relative to the set positions within a given tolerance.

We introduce the following notations of the states: x_1 (drive 1), x_2 (drive 2), x_3 (drive 3), x_4 (computing unit 1), x_5 (computing unit 2), x_6 (power unit 1), x_7 (power unit 2), and x_8 (power unit 3). The output vector consists of $y_1 = x_1$, $y_2 = x_2$, and $y_3 = x_3$. Drive 1 is conventionally divided into two elements x_1 and x_9 to distinguish the operators ORi and ANDi that formalize failure impact at its input $((x_4 \times x_5) + x_7)$.

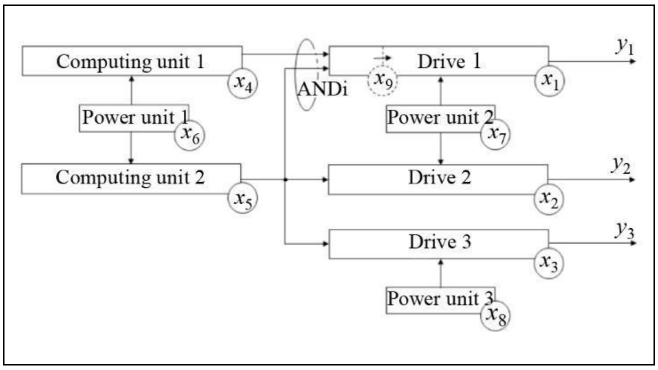

Fig. 4. The functional diagram of the OEC fragment.

The diagram in Fig. 4 is described by the direct logical model (2) of the form

$$\begin{bmatrix} x_1\\ x_2\\ x_3\\ x_4\\ x_5\\ x_6\\ x_7\\ x_8\\ x_9 \end{bmatrix}_{k+1} = \underbrace{\begin{array}{c} \text{ORi}\\ \text{ORi}\\ \text{ORi}\\ \triangledown\\ \triangledown\\ \triangledown\\ \triangledown\\ \triangledown\\ \text{ANDi}\\ \\ \end{array}\begin{array}{c}\begin{bmatrix} \odot & \odot & \odot & \odot & \odot & \odot & 1 & \odot & 1\\ \odot & \odot & \odot & \odot & 1 & \odot & 1 & \odot & \odot\\ \odot & \odot & \odot & \odot & 1 & \odot & \odot & 1 & \odot\\ \odot & \odot & \odot & \odot & \odot & 1 & \odot & \odot & \odot\\ \odot & \odot & \odot & \odot & \odot & 1 & \odot & \odot & \odot\\ \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot\\ \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot\\ \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot\\ \odot & \odot & \odot & 1 & 1 & \odot & \odot & \odot & \odot \end{bmatrix}\\ \begin{array}{ccccccccc} \triangledown & \triangledown & \triangledown & \triangledown & \text{ANDo} & \text{ANDo} & \text{ANDo} & \triangledown & \triangledown \end{array}\end{array}}_{\text{DM}} \vec{\Diamond} \begin{bmatrix} x_1\\ x_2\\ x_3\\ x_4\\ x_5\\ x_6\\ x_7\\ x_8\\ x_9 \end{bmatrix}_k + \begin{bmatrix} x_1\\ x_2\\ x_3\\ x_4\\ x_5\\ x_6\\ x_7\\ x_8\\ x_9 \end{bmatrix}_{\text{init}}, \quad (9)$$

$$\begin{bmatrix} y_1\\ y_2\\ y_3 \end{bmatrix}_k = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}}_{\text{EM}} \times \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & x_9 \end{bmatrix}_k^{\mathrm{T}} = \begin{bmatrix} x_1\\ x_2\\ x_3 \end{bmatrix}_k . \quad (10)$$

The corresponding reverse model (3) is

$$\underbrace{\begin{bmatrix} \hat{x}_1\\ \hat{x}_2\\ \hat{x}_3\\ \hat{x}_4\\ \hat{x}_5\\ \hat{x}_6\\ \hat{x}_7\\ \hat{x}_8\\ \hat{x}_9 \end{bmatrix}_0}_{\hat{X}_0} = \underbrace{\begin{bmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0 \end{bmatrix}}_{\text{EM}^{\mathrm{T}}} \times \underbrace{\begin{bmatrix} y_1\\ y_2\\ y_3 \end{bmatrix}}_{Y_k} + \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0\\ 1 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}}_{\overline{\text{EM}}^{\mathrm{R}}} \times \underbrace{\begin{bmatrix} \&\\ \&\\ \&\\ \&\\ \&\\ \& \end{bmatrix}}_{\mu(\&)} = \begin{bmatrix} y_1\\ y_2\\ y_3\\ \&\\ \&\\ \&\\ \&\\ \&\\ \& \end{bmatrix}, \quad (11)$$

$$\begin{bmatrix} \hat{x}_1\\ \hat{x}_2\\ \hat{x}_3\\ \hat{x}_4\\ \hat{x}_5\\ \hat{x}_6\\ \hat{x}_7\\ \hat{x}_8\\ \hat{x}_9 \end{bmatrix}_{\tau+1} = \underbrace{\begin{array}{c} \triangledown\\ \triangledown\\ \triangledown\\ \triangledown\\ \text{rANDo}\\ \text{rANDo}\\ \text{rANDo}\\ \triangledown\\ \triangledown\\ \\ \end{array}\begin{array}{c}\begin{bmatrix} \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot\\ \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot\\ \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot\\ \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot & 1\\ \odot & 1 & 1 & \odot & \odot & \odot & \odot & \odot & 1\\ \odot & \odot & \odot & 1 & 1 & \odot & \odot & \odot & \odot\\ 1 & 1 & \odot & \odot & \odot & \odot & \odot & \odot & \odot\\ \odot & \odot & 1 & \odot & \odot & \odot & \odot & \odot & \odot\\ 1 & \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot \end{bmatrix}\\ \begin{array}{ccccccccc} \text{rORi} & \text{rORi} & \text{rORi} & \triangledown & \triangledown & \triangledown & \triangledown & \triangledown & \text{rANDi} \end{array}\end{array}}_{\text{rDM}} \overleftarrow{\Diamond} \begin{bmatrix} \hat{x}_1\\ \hat{x}_2\\ \hat{x}_3\\ \hat{x}_4\\ \hat{x}_5\\ \hat{x}_6\\ \hat{x}_7\\ \hat{x}_8\\ \hat{x}_9 \end{bmatrix}_{\tau} + \underbrace{\begin{bmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0 \end{bmatrix}}_{\text{EM}^{\mathrm{T}}} \times \begin{bmatrix} y_1\\ y_2\\ y_3 \end{bmatrix}. \quad (12)$$

Consider the failure of power unit 3 ($x_{8.0}=1$), starting from the direct failure propagation model (9), (10). The corresponding calculations can be performed by the reader independently. At the first step, we obtain $x_{3.1}=1$. At all subsequent steps, $x_{3.k}=1$ and $x_{8.k}=1$. The output vector becomes

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}_k = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \qquad (13)$$

i.e., the failure manifests itself as a malfunction of drive 3.

Now we localize the failure using triplex models.

Step 0. We calculate the initial estimate $\hat{X}_0$ of the failure vector by formula (11):

$$\hat{X}_0 = \begin{bmatrix} 0 & 0 & 1 & \& & \& & \& & \& & \& & \& \end{bmatrix}^{\mathrm{T}}. \qquad (14)$$

According to this estimate, the first and second elements of the logical network (drives 1 and 2) are resistive to failure impact whereas the third element (drive 3) is prone to it (the location of the failure remains unclear), which does not contradict the problem condition. The other elements of the logical network (from the fourth to the ninth) have the indefinite state.

Step 1. We substitute the vector (14) into formula (12). For $\tau = 0$ (see the note to Table 4), the first cycle results in

$$\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \\ \hat{x}_4 \\ \hat{x}_5 \\ \hat{x}_6 \\ \hat{x}_7 \\ \hat{x}_8 \\ \hat{x}_9 \end{bmatrix}_1 = \begin{array}{c} \nabla \\ \nabla \\ \nabla \\ \nabla \\ \text{rANDo} \\ \text{rANDo} \\ \text{rANDo} \\ \nabla \\ \nabla \end{array} \underbrace{\begin{array}{c} \begin{bmatrix} \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot \\ \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot \\ \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot \\ \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot & 1 \\ \odot & 1 & 1 & \odot & \odot & \odot & \odot & \odot & 1 \\ \odot & \odot & \odot & 1 & 1 & \odot & \odot & \odot & \odot \\ 1 & 1 & \odot & \odot & \odot & \odot & \odot & \odot & \odot \\ \odot & \odot & 1 & \odot & \odot & \odot & \odot & \odot & \odot \\ 1 & \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot \end{bmatrix} \\ \begin{array}{ccccccccc} \text{rORi} & \text{rORi} & \text{rORi} & \nabla & \nabla & \nabla & \nabla & \nabla & \text{rANDi} \end{array} \end{array}}_{\text{rDM}} \overleftarrow{\Diamond} \begin{bmatrix} 0 \\ 0 \\ 1 \\ \& \\ \& \\ \& \\ \& \\ \& \\ \& \end{bmatrix}_0 + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$= \begin{array}{c} \nabla \\ \nabla \\ \nabla \\ \nabla \\ \text{rANDo} \\ \text{rANDo} \\ \text{rANDo} \\ \nabla \\ \nabla \end{array} \begin{bmatrix} \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot \\ \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot \\ \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot \\ \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot & \& \times \& \leftarrow \& \\ \odot & 0+0 \leftarrow 0 & \frac{1+\& \leftarrow 1}{\& + 1 \leftarrow 1} & \odot & \odot & \odot & \odot & \odot & \& \times \& \leftarrow \& \\ \odot & \odot & \odot & \& & \& & \odot & \odot & \odot & \odot \\ 0+0 \leftarrow 0 & 0+0 \leftarrow 0 & \odot & \odot & \odot & \odot & \odot & \odot & \odot \\ \odot & \odot & \frac{\& + 1 \leftarrow 1}{1+\& \leftarrow 1} & \odot & \odot & \odot & \odot & \odot & \odot \\ 0+0 \leftarrow 0 & \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot \end{bmatrix}_0 + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

In the scalar notations, using Table 4 (table/row-column), we obtain

$$\hat{x}_1 = 0,$$

$$\hat{x}_2 = 0,$$

$$\hat{x}_3 = 1,$$

$$\hat{x}_4 = \underbrace{\mathbf{\&} \times \&_5 \leftarrow \&}_{\&\ (4/2\text{-}3)} = \&,$$

$$\hat{x}_5 = \underbrace{\underbrace{\underbrace{(\mathbf{0} + 0_7 \leftarrow 0)}_{0\ (4/1\text{-}1)} \times \underbrace{(\frac{\mathbf{1} + \&_8 \leftarrow 1}{\mathbf{\&} + 1_8 \leftarrow 1})}_{1 \text{ или } \&\ (4/1\text{-}2)}}_{0(4/4\text{-}2)} \times \underbrace{(\&_4 \times \mathbf{\&} \leftarrow \&)}_{\&\ (4/1\text{-}3)}}_{0(4/4\text{-}2)} = 0,$$

$$\hat{x}_6 = \underbrace{\& \times \&}_{\&\ (4/4\text{-}3)} = \&,$$

$$\hat{x}_7 = \underbrace{\underbrace{(\mathbf{0} + 0_9 \leftarrow 0)}_{0 \text{ или } \&\ (4/2\text{-}2)} \times \underbrace{(0_5 + \mathbf{0} \leftarrow 0)}_{\& \text{ или } 0\ (4/2\text{-}2)}}_{0\ (4/4\text{-}2)} = 0,$$

$$\hat{x}_8 = \underbrace{\frac{\mathbf{1} + \&_9 \leftarrow 1}{\mathbf{\&} + 1_9 \leftarrow 1}}_{(4/1\text{-}2)} = 1 \text{ или } \&,$$

$$\hat{x}_9 = \underbrace{0_7 + \mathbf{0} \leftarrow 0}_{0\ (4/1\text{-}1)},$$

Here, the possible variants for $\hat{x}_j$ (Table 4) are shown by the fractional line; the values related to $\hat{x}_j$ are set in bold; the values $\hat{x}_q$ appearing in the formulas of Table 4 jointly with $\hat{x}_j$ are indicated by the subscripts *q*. Note that in the formula for $\hat{x}_5$, the multiplication of variants according to cell 4-2 in Table 4 is canceled; in the formula for $\hat{x}_8$, it is preserved.

Thus, after the first cycle, the system state has the estimate

$$\begin{bmatrix}\hat{x}_1\\ \hat{x}_2\\ \hat{x}_3\\ \hat{x}_4\\ \hat{x}_5\\ \hat{x}_6\\ \hat{x}_7\\ \hat{x}_8\\ \hat{x}_9\end{bmatrix}_1 = \begin{bmatrix}0\\ 0\\ 1\\ \&\\ 0\\ \&\\ 0\\ 1 \text{ or } \&\\ 0\end{bmatrix} \begin{array}{l} \text{– no failure,}\\ \text{– no failure,}\\ \text{– failure,}\\ \text{– indefinite state,}\\ \text{– no failure,}\\ \text{– indefinite state,}\\ \text{– no failure,}\\ \text{– indefinite state,}\\ \text{– no failure.}\end{array}$$

The estimation procedure can be continued in different directions due to the ambiguous estimate $\hat{x}_{8.1}$. It is possible to analyze each of the options $\hat{x}_{8.1} = 1$ and $\hat{x}_{8.1} = \&$, thereby increasing the amount of calculations, or to accept $\hat{x}_{8.1} = \&$. The advantage of each option seems unobvious and needs to be studied in a particular case. Let us select the second option.

Then, after the second cycle of the reverse triplex model, we obtain

$$\hat{X}_2 = \begin{bmatrix}0 & 0 & 1 & 0 & 0 & \& & 0 & \& & 0\end{bmatrix}^{\mathrm{T}};$$

after the third cycle, the estimation takes the final form

$$\hat{X}_3 = \begin{bmatrix}0 & 0 & 1 & 0 & 0 & 0 & 0 & \& & 0\end{bmatrix}^{\mathrm{T}},$$

remaining invariable at the subsequent cycles.

Step 2. In this result, one component $\hat{x}_8$ has the indefinite value: it can be either 1 (failure) or 0 (no failure). To clarify the situation, we use the direct logical model (9), substituting alternately $\hat{x}_8 = 1$ and $\hat{x}_8 = 0$. In the first case,

$$\hat{X}'$$

$$= \underbrace{\begin{array}{r|ccccccccc|} \text{ORi} & \odot & \odot & \odot & \odot & \odot & \odot & 1 & \odot & 1\\ \text{ORi} & \odot & \odot & \odot & \odot & 1 & \odot & 1 & \odot & \odot\\ \text{ORi} & \odot & \odot & \odot & \odot & 1 & \odot & \odot & 1 & \odot\\ \nabla & \odot & \odot & \odot & \odot & \odot & 1 & \odot & \odot & \odot\\ \nabla & \odot & \odot & \odot & \odot & \odot & 1 & \odot & \odot & \odot\\ \nabla & \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot\\ \nabla & \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot\\ \nabla & \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot & \odot\\ \text{ANDi} & \odot & \odot & \odot & 1 & 1 & \odot & \odot & \odot & \odot\\ \hline & \nabla & \nabla & \nabla & \nabla & \text{ANDo} & \text{ANDo} & \text{ANDo} & \nabla & \nabla \end{array}}_{\text{DM}}$$

$$\vec{\Diamond}\begin{bmatrix}0\\0\\1\\0\\0\\0\\0\\1\\0\end{bmatrix}_0 + \begin{bmatrix}0\\0\\1\\0\\0\\0\\0\\1\\0\end{bmatrix}_0 = \begin{bmatrix}0\\0\\1\\0\\0\\0\\0\\1\\0\end{bmatrix}.$$

The reader can verify that the second case gives $\hat{X}'' = \begin{bmatrix}0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0\end{bmatrix}^{\mathrm{T}}$. In both cases, there is an observation of failure (10), (13), which confirms the possible failures of drive 3 (x_3) and power unit 3 (x_8) in the first case and the possible failure of drive 3 (x_3) in the second case.

This methodical example is simple, and the cause of the failure of drive 3 (x_3) can be established by the reader independently. Two variants are possible here: either drive 3 (x_3) has failed under the indefinite state of power unit 3 (x_8) or power unit 3 (x_8) has failed, causing drive 3 (x_3) to fail as well.

Thus, the modeling result obtained in this example does not contradict the engineering analysis.

CONCLUSIONS

This paper has considered three solutions to monitor the technical condition of components in reconfigurable redundant OECs. The choice of an appropriate variant depends on different factors, including the level of theoretical and applied development, the goals and capabilities of the OEC developer, the criticality of systems under diagnosis, etc.

The most accessible solution is using BiC in the nearest perspective. The next level implies the supplementary application of logical pair monitoring to increase the reliability of diagnosis results significantly under the inevitable errors of diagnostic means. In the distant perspective, it seems reasonable to add algorithms based on the logical failure impact propagation models of OECs.

The algorithms with logical (triplex) models have several features that make the approach attractive:

- Due to their exceptional simplicity, logical models can be effectively applied even for very complex OEC architectures.
- Model building is based on FMEA, a well-mastered methodology for the aircraft industry with acceptable depth and breadth of coverage for the operation conditions of OECs.
- Triplex models are handled using special constructs similar to matrix ones; appropriate methods and software tools have to be developed.

Further research will focus on the analytical determination of the failure estimate vector to reduce the number of iterations when identifying the operable state of OECs.

Acknowledgments. *The authors are grateful to A.M. Ageev for his invaluable contribution to separate aspects of the proposed approach.*

REFERENCES

1. Zhang, Y. and Jiang, J., Bibliographical Review on Reconfigurable Fault-Tolerant Control Systems, *Proc. the 5th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes*, Washington, DC, 2003, pp. 265–276.
2. Willsky, A.S., A Survey of Design Methods for Failure Detection in Dynamic Systems, *Automatica*, 1976, no. 12, pp. 601–611.
3. Patton, R.J., Frank, P.M., and Clark, R.N., *Issues of Fault Diagnosis for Dynamic Systems*, London: Springer, 2000.
4. Bukirev, A.S., Savchenko, A.Y., Yatsechko, M.I., and Malyshev, V.A., Diagnostic System for the Technical Condition of the Aircraft Avionics Complex Based on Intelligent Information Technologies, *Modeling, Optimization and Information Technology*, 2020, vol. 8, no. 1(28), pp. 12–24. (In Russian.)
5. *Diagnosis and Fault-tolerant Control. Vol. 1: Data-driven and Model-based Fault Diagnosis Techniques*, coordinated by V. Puig and S. Simani, London: ISTE Ltd; Hoboken: John Wiley & Sons, 2021.
6. *Diagnosis and Fault-tolerant Control. Vol. 2: From Fault Diagnosis to Fault-tolerant Control*, coordinated by V. Puig and S. Simani, London: ISTE Ltd; Hoboken: John Wiley & Sons, 2021.
7. Zhao, R., Yan, R., Chen, Z., et al., Deep Learning and Its Applications to Machine Health Monitoring, *Mechanical Systems and Signal Processing*, 2019, no. 115, pp. 213–237.
8. Parkhomenko, P.P. and Sogomonyan, E.S., *Osnovy tekhnicheskoi diagnostiki* (*optimizatsiya algoritmov diagnostirovaniya, apparatnye sredstva*) (Foundations of Technical Diagnosis (Optimization of Diagnostic Algorithms, Hardware)), Moscow: Energoatomizdat, 1981. (In Russian.)
9. Sapozhnikov, V.V., Sapozhnikov, Vl.V., and Efanov, D.V., Synthesis of a Built-in Control Circuit for Multi-Output Combinational Devices Based on Logical Complement and Signal Compression, *Journal of Instrument Engineering*, 2020, vol. 63, no. 7, pp. 583–599. (In Russian.)
10. Ageev, A.M., Principles of Storing and Monitoring Configuration Information in the Task of On-Board Equipment Complex Redundancy Managing, *Mechatronics, Automation, Control*, 2022, vol. 23, no. 1, pp. 45–55. (In Russian.)
11. Chandler, P.R., Self-repairing Flight Control System Reliability and Maintainability Program Executive Overview, *Proc. IEEE National Aerospace and Electronics Conf.*, Dayton, 1984, pp. 586–590.
12. Nishiyama, T., Suzuki, Sh., Sato, M., and Masui, K., Simple Adaptive Control with PID for MIMO Fault Tolerant Flight Control Design, *AIAA*, 2016, art. no. 0132.
13. Mel'nik, E.V., Methods and Software Tools for Increasing the Reliability of Network Information and Control Systems Based on Computer Resource Reconfiguration, *Doctoral (Eng.) Dissertation*, Taganrog: Research Institute of Multiprocessor Computing Systems, Southern Federal University, 2014. (In Russian.)
14. Zaets, V.F., Abdulin, R.R., Kulabuzov, V.S., et al., RF Patent 2629454 S2, *Byull. Izobret.*, 2017, no. 8.
15. Bukov, V.N., Ageev, A.M., Evgenov, A.V., and Shurman, V.A., *Upravlenie izbytochnost'yu tekhnicheskikh sistem. Supervizornyi sposob upravleniya konfiguratsiyami* (Redundancy Management of Technical Systems. The Supervision Method of Configuration Management), Moscow: INFRA-M, 2023. (In Russian.)
16. *GOST* (State Standard) *R 55255–2012: Air Transport. Maintenance and Repair System for Aircraft Equipment. Organization of Works to Diagnose the Technical Condition of Aircraft Equipment*, Moscow: Standartinform, 2020.
17. Bolelov, E.A., Matyukhin, K.N., Prokhorov, A.V., and Prokof'ev, I.O., *Tekhnicheskie sredstva kontrolya pri ekspluatatsii radioelektronnogo oborudovaniya vozdushnogo transporta* (Technical Means of Control in the Operation of Radio-Electronic Equipment of Air Transport), Moscow: Zhukovsky Academy, 2018. (In Russian.)
18. Bukov, V.N., Ozerov, E.V., and Shurman, V.A., Pair Monitoring of Redundant Technical Systems, *Automation and Remote Control*, 2020, vol. 81, no. 1, pp. 74–93.
19. Bukov, V.N., Bronikov, A.M., and Sel'vesyuk, N.I., An Algorithm to Localize Onboard Complex Failures Based on Mixed Directed Graphs, *Problemy Bezopasnosti Poletov*, 2010, no. 2, pp. 57–71. (In Russian.)
20. *Diagnostirovanie i prognozirovanie tekhnicheskogo sostoyaniya aviatsionnogo oborudovaniya* (Technical Condition Diagnosis and Forecasting for Aircraft Equipment), Sindeev, I.M., Ed., Moscow: Transport, 1984. (In Russian.)
21. Dzhandzhgava, G.I., Dyadischev, A.V., and Garifov, R.Sh., On a Concept for Technical Condition Monitoring of Methods

Used for Analyzing Physical Media, *Ideas and Innovations*, 2018, vol. 6, no. 3, pp. 64–68. (In Russian.)
22. Mozgalevskii, A.V. and Kalyavin, V.P., *Sistemy diagnostirovaniya sudovogo oborudovaniya* (Diagnosis Systems for Shipboard Equipment), Leningrad: Sudostroenie, 1987. (In Russian.)
23. Sokolov, N.L., The Basic Principles of Diagnostics of Serviceability of the Onboard Equipment Automatic (SV) and Development of the Recommendations on Elimination of Not Regular Situations, *Advances in Current Natural Sciences*, 2007, no. 6, pp. 16–20. (In Russian.)
24. Baranovsky, A.M. and Privalov, A.E., Onboard Monitoring and Diagnostic System of Small Space Vehicles, *Journal of Instrument Engineering*, 2009, vol. 52, no. 4, pp. 51–65. (In Russian.)
25. Bukov, V.N., Ozerov, E.V., and Shurman, V.A., Logical Pair Monitoring That Accounts for the Grey Zone, *Automation and Remote Control*, 2020, vol. 81, no. 6, pp. 1037–1050.
26. Bukov V.N., Aver'yanov I.N., Bronnikov A.M., et al., RF Patent 2557441 S2, *Byull. Izobret.*, 2015, no. 7.
27. *GOST* (State Standard) *R 27.606-2013: Reliability in Engineering. Reliability Management. Safety-Oriented Maintenance*, Moscow: Standartinform, 2014.
28. Bukov, V.N., *Vlozhenie sistem. Analiticheskii podkhod k analizu i sintezu matrichnykh sistem* (Nested Systems. An Analytical Approach to the Analysis and Design of Matrix Systems), Kaluga: Bochkareva's Press, 2006. (In Russian.)
29. Bukov, V.N., Bronnikov, A.M., and Sel'vesyuk, N.I., A Failure Propagation Model for the Helicopter Altitude-Velocity Channel, *Problemy Bezopasnosti Poletov*, 2010, no. 10, pp. 39–51. (In Russian.)
30. Bronnikov, A.M. and Morozov, D.V., Troubleshooting of Directly not Observable Failures of Airborne Systems Based on Mixed Directed Graph, *Mekhatronika, Avtomatizatsiya, Upravlenie*, 2013, no. 1, pp. 62–66. (In Russian.)

This paper was recommended for publication by V.G. Lebedev, a member of the Editorial Board.

Received June 11, 2023, and revised September 24, 2023. Accepted September 28, 2023.

Author information

Bukov, Valentin Nikolaevich. Dr. Sci. (Eng.), Institute of Aircraft Equipment, Zhukovsky, Russia
✉ v_bukov@mail.ru
ORCID iD: https://orcid.org/0000-0002-5194-8251

Bronnikov, Andrei Mikhailovich. Dr. Sci. (Eng.), Bauman Moscow State Technical University, Moscow, Russia
✉ bronnikov_a_m@mail.ru

Vorob'ev, Aleksandr Vladimirovich. Dr. Sci. (Eng.), Institute of Aircraft Equipment, Zhukovsky, Russia
✉ vorobiev@niiao.ru

Popov, Aleksandr Sergeevich. Cand. Sci. (Eng.), Zhukovsky–Gagarin Air Force Academy, Voronezh, Russia
✉ saga30@yandex.ru

Shurman, Vladimir Aleksandrovich. Ramenskoe Instrument-Making Design Bureau, Zhukovsky Branch, Zhukovsky, Russia
✉ vshurman@rpkb.ru

Cite this paper
Bukov, V.N., Bronnikov, A.M., Vorob'ev, A.V., Popov, A.S., and Shurman, V.A., Component Monitoring to Manage the Redundancy of an Onboard Equipment Complex. *Control Sciences* **5,** 75–91 (2023). http://doi.org/10.25728/cs.2023.5.7

Translated into English by *Alexander Yu. Mazurov*, Cand. Sci. (Phys.–Math.), Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
✉ alexander.mazurov08@gmail.com