Таврический Вестник Информатики и Математики

№ 4 (33) ' 2016

МЕЖДУНАРОДНОЕ НАУЧНО-ТЕОРЕТИЧЕСКОЕ ИЗДАНИЕ
КРЫМСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ им. В. И. ВЕРНАДСКОГО

ОСНОВАН В 2002 ГОДУ

Свидетельство о регистрации средства массовой информации

ПИ №ФС77-61826 от 18.05.2015

TAURIDA JOURNAL OF COMPUTER SCIENCE THEORY AND MATHEMATICS

2016, No. 4

INTERNATIONAL THEORETICAL RESEARCH EDITION, PEER–REVIEWED JOURNAL
CRIMEA FEDERAL V. I. VERNADSKY UNIVERSITY

FOUNDED IN 2002
The State Registration Certificate
ПИ No ФС77-61826 (18.05.2015)

ISSN 1729-3901
Key title: Tavriceskij vestnik informatiki i matematiki

АДРЕС РЕДАКЦИИ:

Крымский федеральный университет им. В. И. Вернадского
пр-т Вернадского, 4, г. Симферополь, Республика Крым, Российская Федерация, 295007

ДЛЯ ПЕРЕПИСКИ:

Таврический Вестник Информатики и Математики
Факультет математики и информатики, Таврическая академия КФУ им. В.И. Вернадского
пр-т Вернадского, 4, г. Симферополь, Республика Крым, Российская Федерация, 295007

Тел. гл. редактора:	+7 (3652) 63-75-42
Тел. редакции:	+7 (3652) 602-466
e-mail (гл. редактор):	vidonskoy@mail.ru
e-mail (для переписки):	article@tvim.info
сайт журнала:	www.tvim.info

**Журнал публикует оригинальные и обзорные статьи
по вопросам теоретической и прикладной информатики и математики**

Ведущие тематические разделы:

Функциональный анализ и его приложения

Интегральные, дифференциальные уравнения и динамические системы

Спектральные и эволюционные задачи

Математические проблемы гидродинамики

Дискретная оптимизация

Математическая логика, теория алгоритмов и теория сложности вычислений

Машинное обучение, распознавание и извлечение закономерностей

Дедуктивные системы и базы знаний

Знаниеориентированные и гибридные математические модели принятия решений

Синтез моделей принятия решений при неполной начальной информации

Вычислительная математика

Печатается по решению ученого совета Крымского федерального университета имени В. И. Вернадского.
Протокол № 11 от 30 ноября 2016 г.

OFFICE ADDRESS:

V. I. Vernadsky Crimean Federal University
Vernadsky Avenue, 4, Simpferopol, Crimea, 295007, Russian Federation

JOURNAL SITE: www.tvim.info

FOR CORRESPONDENCE:

Taurida Journal of Computer Science Theory and Mathematics,
Faculty of Mathematics and Computer Science Theory, Taurida Academy of Vemadsky CFU
Vernadsky Avenue, 4, Simpferopol, Crimea, 295007, Russian Federation

Tel. +7 3652 637 542 — editor-in-chief
+7 3652 602 466 — office
Email: vidonskoy@mail.ru — editor-in-chief
article@tvim.info — for correspondence

Taurida Journal of Computer Science Theory and Mathematics is a peer-reviewed journal, published by V.I. Vernadsky Crimean Federal University. The journal publishes research papers in the fields of computer science and mathematics.

THEMATIC SECTIONS:

Algorithm Theory, Mathematical Logic, Discrete Optimization, Complexity Theory, Calculus Mathematics, Machine Learning, Pattern Recognition, Data Mining, Deductive Systems and Knowledge Bases, Decision Making Models;

Functional Analysis and Applications, Integral, Differential Equations, Dynamic Systems, Spectral and Evolutional Tasks, The mathematical problems of hydrodynamics.

СОДЕРЖАНИЕ

TABLE OF CONTENTS

UDC: 519.816.4

MSC2010: 90C29

RISKS IN A MULTICRITERIA PROBLEM UNDER UNCERTAINTY

© Zhukovskiy V.I., Kirichenko M. M., Boldyrev M. V.

Moscow State University named after Lomonosov
Faculty of Computational Mathematics and Cybernetics
Department of Optimal Control
Leninskiye Gory, GSP-1, Moscow, 119991, Russian Federation
e-mail: *zhkvlad@yandex.ru, moyomylo11@gmail.com, boldyrev@list.ru*

© Smirnova L. V.

State University of Humanities and Technology
Computer Sciences Department
142611, Moscow Region, Orekhovo-Zuevo, Zelyonaya Street, 22
e-mail: *smirnovalidiya@rambler.ru*

Abstract. Applicability and novelty of the following research lies in that the decision-maker in a multicriteria problem aims not only to maximize guaranteed values of each criterion, but also to minimize the guaranteed risks accompanying the said maximization. The topic of the research lies at the interface of the multicriteria problem (MP) theory and the Savage–Nichans minimax regret principle (MRP): the concept of a weakly effective estimate has been derived from the MP theory, while estimation of risks with values of the Savage–Nichans regret function has been derived from the MRP. The scope of this research is limited to interval uncertainties: the decision-maker only knows the limits of the interval, and probabilistic characteristics are missing. A new term of a "strongly guaranteed solution under outcomes and risks" (SGSOR) is introduced; its existence for "regular"–confined–strategies for the mathematical programming is established. As an example of a practical application, a problem of diversification of a multi-currency deposit has been suggested and solved.

Keywords: *Pareto maximum, strategy, uncertainty, vector guarantee, Savage risk, principle of minimax regret.*

Introduction

A multicriteria problem under uncertainty (MPU) is defined here by an ordered set as follows:

$$< \mathbb{N}, X, Y, f(x,y) >, \tag{1}$$

where $\mathbb{N} = 1, \dots, N$ ($N \geq 2$) is a set of index numbers of the criteria; possible strategies $x \in X \subseteq \mathbb{R}^n$ (belong to the set X from Euclidian n-dimensional space $\mathbb{R}^n$); $Y \subseteq \mathbb{R}^m$ is a subset of interval uncertainties y; $f(x,y)$ is a vector of criteria $f_i(x,y)$ ($i \in \mathbb{N}$); each criterion $f_i(x,y)$ is a scalar function, defined in the Cartesian product $X \times Y$. When building a Savage–Nichans risk function, so-called strategic uncertainties $y(x) : X \to Y$ are used; sets of such uncertainties are denoted as Y^X. Therefore, the decision-maker in the problem (1) seeks to adopt a strategy $x \in X$ to increase values of each criterion $f_i(x, y)$ (called *outcomes*). The decision maker also must take in account the effects from any possible uncertainty $y \in Y$. In the present article, we will additionally consider new N criteria: Savage–Nichans risk functions. Thus the present article is dedicated to mathematical rationalization of a method of building a strategy that would both increase guaranties of all outcomes and decrease risks accompanying such an increase.

1. Savage–Nichans minimax regret

In 1939, Romanian mathematician, who migrated to America in 1938, Abraham Wald (1902–1950) introduced [1] the maximin principle (principle of guaranteed outcome), which allows finding a guaranteed result in a one-criterion problem under uncertainty (OCPU). In almost 10 years, German mathematician Jürg Nichans in 1948 and American mathematician, economist, and statistician Leonard Savage (1917–1971) in 1951 suggested [2, 3] the principle of minimax regret, which allows building guaranteed risks for OCPU; it would later be referred to as "Savage risk" (later named "Nichans–Savage criterion"). We will note that Savage worked as a statistician under the commandership of John von Neumann, which motivated the emergence of the minimax regret principle. In 1977, the Savage Award was founded; it is annually conferred to authors of the two most outstanding doctoral theses in the fields of economics and statistics. MRP for a one-criterion problem

$$\Gamma_1 = \langle X, Y, \phi(x, y) \rangle ,$$

centered around building a pair $(x^r, R^r_\phi) \in X \times \mathbb{R}$ compliant with the following equations:

$$R^r_\phi = \max_{y \in Y} R_\phi(x^r, y) = \min_{x \in X} \max_{y \in Y} R_\phi(x, y), \tag{2}$$

where Savage–Nichans risk function is:

$$R_\phi(x, y) = \max_{z \in X} \phi(z, y) - \phi(x, y). \tag{3}$$

The value $R_\phi(x, y)$ from (3) is called (Savage–Nichans) *risk* in Γ_1. The risk function $R_\phi(x, y)$ estimates the margin between the value of the criterion and the "best" (for the decision-maker in the problem Γ_1) value $\max_{z \in X} \phi(z, y)$. Obviously, the decision-maker tries to minimize $R_\phi(x, y)$ by selecting their strategy $x \in X$ assuming the greatest possible resistance in accordance with the principle of guaranteed result. Therefore, following (2)–(3), the decision maker is *optimistic* (aims for "the best" value $\max_{x \in X} \phi(x, y)$) as opposed to *pessimistic* decision-maker (who is oriented towards "the worst" solution: the maximin solution $(x^0, \phi^0 = \max_{x \in X} \min_{y \in Y} \phi(x, y) = \min_{y \in Y} \phi(x^0, y))$. From now on, we will consider that the decision-maker in the problem (1) exposes an optimistic strategy: they make a Savage–Nichans risk function

$$R_i(x, y) = \max_{z \in X} f_i(z, y) - f_i(x, y), \qquad i \in \mathbb{N} \tag{4}$$

for each component of the vectorial criterion $f(x, y)$; its value is called *risk*. We will note two conditions: first, each criterion $f_i(x, y)$ from (1) corresponds to its own risk $R_i(x, y)$. Second, the decision-maker aims to choose a strategy $x \in X$ to decrease all risks $R_i(x, y)$ simultaneously.

Remark 1. Such models can be found, for example, a market, where a salesperson under conditions of unpredictable import tries to maximize their profit. Review of these uncertainties can be found in [4, p. 102–106], [5, p. 19–32] and other publications.

Existence of uncertainties forms the set

$$\phi(x, Y) = \{\phi(x, y) | \forall y \in Y\}$$

for all $x \in X$. The set can be "narrowed down" using risks. What is a risk? Famous Russian optimization theory expert T.K. Sirazetdinov believes there is no strict definition of risk [6, p. 31]. Sixteen definitions have been reviewed in monography [7, p. 15]. A majority of these requires statistical data on uncertainty. However, the decision-maker often does not have such data. Such situations fall in the scope of the present article.

We will define *risk* as *possibility of deviance of the values from the desired ones*. We will note that this definition echoes the "regular" microeconomic risks described in, for example, [8, pp. 40–50]; it corresponds to the aforesaid risk. Lastly, accounting for risks is an actual problem in economics; this notion is supported by the 1990 Nobel Prize in economics, won by Harry Max Markowitz and his new approach [9] to the analysis of investment distribution risk and diversification of the expected return on investment. In the present article, this approach is extended to multicriteria problems. Finally, in publications on macroeconomics [4, p. 103], [8, p. 5], [10, p. 343] decision-makers are divided into three categories: risk-averse, risk-loving, and risk-neutral. In the present article, the decision-maker is considered risk-neutral (and, as noted above, optimistic).

2. Strong guaranties and transition from (1) to a $2N$–criteria problem

By forming a risk function $R_i(x, y)$ for each criterion $f_i(x, y)$ using formula (4), we transform the original problem into a following extension of the problem (1) to a $2N$–criteria problem

$$\langle \mathbb{N}, X, Y, \{f_i(x, y), -R_i(x, y)\}_{i\in\mathbb{N}} \rangle \,, \tag{5}$$

where $\mathbb{N}, X$, and Y are identical to those in (1), and the vectorial criterion $f(x, y)$ is concatenated with a N–vector $-R(x, y) = (-R_1(x, y), ..., -R_N(x, y))$, where the latter are deliberately taken in negation to provide uniformity of applying strategies $x \in X$ to each of the $2N$criteria $f_i(x, y), -R_i(x, y)$. Precisely, in (5), the decision-maker forms their strategy $x \in X$ in order to increase value of each criterion $f_i(x, y), -R_i(x, y)(i \in \mathbb{N})$ from both N–vectors: $f(x, y)$ and $-R(x, y)$. The decision-maker must also expect a possibility of emergence of uncertainty $y \in Y$.

We will now proceed to the strong guaranties for criteria. The first author of the present article suggested three methods of accounting for uncertainties in the series of works [11], [12]: analog of a saddle point [11] and two analogs of maximin [12]: strong and vectorial guarantees. It should be noted that the former guaranty is used in the present work, while the latter was used in [13].

Definition 1. A scalar function $f_i[x]$ is called a **strong guaranty** of criterion $f_i(x, y) : X \times Y \to \mathbb{R}$, if for each $x \in X$ the in inequality takes plays

$$f_i[x] \leq f_i(x, y)], \qquad \forall y \in Y \quad (i \in \mathbb{N}).$$

Remark 2. Obviously, the function $f_i[x] = \min_{y\in Y} f_i(x, y)$ $\forall x \in X$ is a strong guaranty for $f_i(x, y)$. This points us to a method of building strong guaranties for the $2N$ criteria from (5).

We will now consider strong guaranties $R_i[x]$ for risk functions $R_i(x, y)$ from (4). Building such guaranties is conducted in three stages.

First, function

$$\psi_i(y) = \max_{z\in X} f_i(z, y) \qquad \forall y \in Y \quad (i \in \mathbb{N}),$$

is defined.

Second, Savage–Nichans risk functions are built:

$$R_i(x, y) = \psi_i(y) - f_i(x, y) \qquad (i \in \mathbb{N}).$$

Third, the strong guaranty is built:

$$\mathbb{R}_i[x] = \max_{y \in Y} R_i(x, y) \qquad (i \in \mathbb{N}).$$

We'll note that the decision-maker seeks to decrease their risk $R_i(x, y)$ by choosing a suitable $x \in X$ for a given realization of $\forall y \in Y$. If the said functions $f_i[x], -R_i[x]$ exist, then they are strong guaranties for $f_i(x, y), -R_i(x, y)$, correspondingly. Indeed,

$$f_i[x] = \min_{y \in Y} f_i(x, y) \implies f_i[x] \le f_i(x, y) \qquad \forall y \in Y \quad (i \in \mathbb{N}),$$

$$-R_i[x] = \min_{y \in Y} (-R_i(x, y)) \implies -R_i[x] \le -R_i(x, y) \qquad \forall y \in Y \quad (i \in \mathbb{N}).$$

Existence of $f_i[x]$ and $R_i[x]$ follows from the well-known fact of operations research:

Lemma 1. [14, p. 54] *If sets X and Y are compact and criteria $f_i(x, y)$ are continuous in $X \times Y$, then*
a) function $f_i[x] = \min_{y \in Y} f_i(x, y)$ is continuous in X;
b) function $\psi_i[x] = \max_{x \in X} f_i(x, y)$ is continuous in X.

From this point, *comp* $\mathbb{R}^n$ is the designation for the set of compacts within $\mathbb{R}^n$, and the continuousness of $f_i(x, y)$ in $X \times Y$ is denoted as $f_i(x, y) \in C(X \times Y)$.

Remark 3. In problem (1), if criterion $f_i(x, y) \in C(X \times Y)$ and $X \in comp\,\mathbb{R}^n$, $Y \in comp\ \mathbb{R}^m$, then the Savage risk function $R_i(x, y)$ $(i \in \mathbb{N})$ in (4) is continuous in $X \times Y$. Indeed, continuousness of $\phi_i[y]$ follows from Lemma 1, statement a), and, accordingly, the Savage risk function is continuous as a difference of continuous functions: $R_i(x, y) = \phi_i[y] - f_i(x, y)$ $(i \in \mathbb{N})$.

Remark 4. The Savage risk function from (4) characterizes deviation of criterion $f_i(x, y)$ from the "desired", value $\max_{x \in X} f_i(x, y)$. This prompts the decision maker to choose a strategy $x \in X$, which would decrease the difference $R_i(x, y)$ from (4) to the greatest extent possible or, equivalently, increase $-R_i(x, y)$. Then the given MPU (5) is assigned to a $2N$-criteria problem

$$\langle X, Y, \{f_i(x, y), -R_i(x, y)\}_{i \in \mathbb{N}} \rangle \,.$$

The goal of the decision maker in (5) is to choose a strategy $x \in X$ so that the $2N$ criteria $f_i(x, y), -R_i(x, y)$ $(i \in \mathbb{N})$ assume the greatest values possible; the decision maker must also expect a possibility of emergence of uncertainty $y \in Y$.

3. Formalization of a guaranteed under outcomes and risks solution of the problem (1)

Literature dedicated to MPUs is quite extensive (in particular, we refer the reader to [13]). The distinctiveness of the "interval character" of uncertainties in (1) (about which we know only its limits) prompts us to refer to available information: the decision-maker learns guaranteed changes of criteria. In the present work, we limit ourselves to strong guaranties $f_i[x]$ and $-R_i[x]$ of criteria $f_i(x,y)$ and $-R_i(x,y)$. Therefore we find natural transition from MPU (5) to a "multicriteria problem of guaranties", in which uncertainty does not exist:

$$\Gamma^g = \left\langle X, \{f_i[x], -R_i[x]\}_{i\in\mathbb{N}} \right\rangle .$$

Finally, criteria $f_i[x]$ and $-R_i[x]$ in $\Gamma^g$are closely linked by an optimization of their "semantics," since $R_i[x]$ assesses the risk of the formation of the outcome $f_i[x]$: an increase in the difference $f_i[x] - R_i[x]$ entails an increase of the guaranteed equal outcome $f_i[x]$, or a decrease of the risk of $R_i[x]$. Inversely, a decrease in this difference involves an increase $f_i[x]$ or an increase in the risk of f $R_i[x]$. The goal of the decision-maker is to increase $f_i[x]$, or a decrease of the risk of $R_i[x]$ at the same time a possible increase $f_i[x]$ involves decrease of the risk of $R_i[x]$. Thus we will link the $2N$-criteria problem Γ^g to an auxiliary N-criteria problem:

$$\Gamma^a = \left\langle X, \{F_i[x] = f_i[x] - R_i[x]\}_{i\in\mathbb{N}} \right\rangle . \tag{6}$$

To formalize a solution of (1) with regards to guaranteed outcomes and risks, we will use one of the definitions of vectorial maximum used in the multicriteria problem theory [15]. First such definition was formulated in 1909 by Italian economist and sociologist Vilfredo Pareto (1848–1923) and, therefore, was eventually named the "Pareto maximum" (effective solution).

We will, however, use the concept of "Slater maximality", which includes Pareto maximality as a particular case. This concept apparently first emerged in Russian literature in article of L. Hurwitz translated into Russian [16].

Definition 2. A strategy x^S is called *Slater maximal* (sometimes *weakly effective*) for a N-criteria problem (6) if $\forall x \in X$ the following system is incompatible:

$$F_i[x] > F_i[x^S] \qquad (i \in \mathbb{N}).$$

Remark 5. From Definition 2 follows that $x^* \in X$ is not Slater maximal in (6), if $\exists \bar{x} \in X$ the following inequalities are correct:

$$F_i[\bar{x}] > F_i[x^*] \qquad (i \in \mathbb{N}).$$

Statement 1. *(sufficient condition). If*

$$\min_{i\in\mathbb{N}} F_i[x^S] = \max_{x\in X} \min_{i\in\mathbb{N}} F_i[x], \tag{7}$$

then x^S is Slater maximal in (6).

Proof. From (7) and Remark 5 follows: $[\forall x \in X \;\; \exists j \in \mathbb{N} : F_j[x] \le F_j[x^S]] \implies$ [system $\left\{F_i[x] \le F_i[x^S]\right\}_{i\in\mathbb{N}}$ is incompatible] $\implies$ [x^S is Slater maximal in (6)]. □

Theorem 1. *(existence). If $f_i(\cdot) \in C(X \times Y)$ $i \in \mathbb{N}$ and X and Y are compact, then there is a Slater maximal strategy in* (6).

Proof. From Lemma 1 follows:

$$f_i(\cdot) \in C(X \times Y) \implies f_i[x] \in C(X) \qquad \forall (i \in \mathbb{N}),$$

and from Remark 3: $R_i(\cdot) \in C(X \times Y)$ $(i \in \mathbb{N})$. Then Lemma 1 provides us with the following: $\min_{i\in\mathbb{N}} F_i[x] = \min_{i\in\mathbb{N}}(f_i[x] - R_i[x]) \in C(X)$ $(i \in \mathbb{N})$. Continous functions $F_i[x]$ reach their maximum in some point x^S, where (7) is correct. The statement being proved follows from Statement 1.

□

Definition 3. A triplet $(x^S, f[x^S], R[x^S])$ is called *strongly guaranteed under risks and outcomes* solution of (1) if there exist

$$f_i[x] = \min_{y\in Y} f_i(x,y), R_i[x] = \max_{y\in Y} R_i(x,y) \qquad (i \in \mathbb{N})$$

and x^S the Slater maximum for the problem (6).

We will remind the reader that

$$f[x] = (f_1[x], \dots, f_N[x]), \;\; R[x] = (R_1[x], \dots, R_N[x]),$$
$$R_i[x] = \max_{y\in Y} R_i(x,y), \;\; R_i(x,y) = \max_{z\in X} f_i(z,y) - f_i(x,y).$$

Why is the strongly guaranteed under outcomes and risks solution (SGOR) suggested as a “good” solution of the MPU (1)?

First, it provides an answer to the indigenous Russian question: “What is to be done?” It is suggested the decision maker follow the strategy x^S of the tripod $(x^S, f[x^S], R[x^S])$.

Second, this strategy x^S “provides” outcomes $f_i(x^S, y)$ greater than or equal to $f_i[x^S]$ with a risk of $R_i(x^S, y)$ not greater than $R_i[x^S]$ for any $i \in \mathbb{N}$, under implementation of any uncertainty $y \in Y$ (i.e. x^S set lower limits for outcomes implemented under $x = x^S$ and upper limits for risks accompanying this implementation).

Third, situation x^S implements “the greatest” Slater maximal outcomes and corresponding “minus” risks; otherwise speaking, there is no situation $x \neq x^S$, in which all risk guarantees $f_i[x^S]$ would increase and at the same time, all risk guarantees would decrease $R_i[x^S]$.

It should be noted that the union of “second,” and “third”, is somewhat analogous to the maximin strategy for a one-criterion problem under uncertainty, with the difference being the substitution of the inner minimum in Γ_1 with $\min_{y \in Y} F_i(x, y)$ $(i \in \mathbb{N})$ and the substitution of the outer maximum with the Slater maximum. Two possible research routes are available at this point. The first one of them is to substitute the Slater maximum with the Slater, Borwein, Geoffrion optimums and conical optimality, and to establish connection between such different solutions. The second route is based on the desire of the decision maker for higher profits, while the guarantees in Definition 2 are the “lowest”. Therefore it is possible to substitute scalar minimums (from the inner minimum in the maximin) with a vectorial one (listed above), therefore increasing the guarantees. Connections between solutions are also of interest (some attempts to build such connection the first author lists in the monography [13]).

Fourth, Definition 2 allows suggesting a constructive method of building a SGOR. It may be reduced to four steps.

Step 1. Using $f_i(x, y)$, find $\max_{x \in X} f_i(x, y) = \psi_i[y]$ and build the Savage risk function for criterion $f_i(x, y)$, assuming $R_i(x, y) = \psi_i[y] - f_i(x, y) (i \in \mathbb{N})$.

Step 2. Build outcome guarantees $f_i[x] = \min_{y \in Y} f_i(x, y)$ and risk guarantees $R_i[x] = \max_{y \in Y} R_i(x, y)$ $(i \in \mathbb{N})$.

Step 3. For a N-criteria guarantees problem $\left\langle X, Y, \{F_i[x] = f_i[x] - R_i[x]\}_{i \in \mathbb{N}} \right\rangle$, calculate a Slater maximal strategy x^S. Remark 1 may be used here as well as the following transition to the concept of Pareto maximality.

Definition 4. Strategy $x^P \in X$ is called **Pareto maximal** (sometimes also **effective**) for a N-criteria problem (6), if for any $x \in X$, the system of inequalities

$$F_i[x] \geq F_i[x^P] \qquad (i \in \mathbb{N}),$$

of which at least one inequality is strict, is incompatible.

We will note the following:

First, as follows from Definitions 2 and 3, any Pareto maximal strategy is at the same time Slater maximal; the opposite is not generally correct.

Second, according to the Karlin's lemma [6, p. 71] the strategy $x^P \in X$ satisfying the following condition:

$$\max_{x\in X}\sum_{i\in\mathbb{N}}\alpha_i F_i[x] = \sum_{i\in\mathbb{N}}\alpha_i F_i[x^P],$$

where α_i = const > 0, is Pareto maximal for the problem (6) Assuming constant $\alpha_1 = \alpha_2 = 1$, we obtain the following equation, suitable for building Pareto (and thus Slater) maximums:

$$\max_{x\in X}(F_1[x] + F_2[x]) = F_1[x^S] + F_2[x^S].$$

Step 4. Using x^S, determine the values of guarantees $f_i[x^S]$ and $R_i[x^S]$ $(i \in \mathbb{N})$ and collect them into two N-vectors $f[x^S] = (f_1[x^S], \dots, f_N[x^S]), R[x^S] = (R_1[x^S], \dots, R_N[x^S])$.

The resulting tripod $(x^S, f[x^S], R[x^S])$ forms the desired SGOR, which complies with Definition 2, i.e. using strategy x^S results in a guaranteed outcome $f_i[x^S]$ with a guaranteed Savage risk $R_i[x^S]$ for each criterion $f_i(x, y)$ $(i \in \mathbb{N})$.

4. Risks and outcomes for diversification of a deposit into sub-deposits in different currencies

As we mentioned before, there are three types of decision-makers: risk-averse, risk-loving, and risk-neutral. In the present article, a solution for the problem of diversification of a deposit into sub-deposits in rubles and in a foreign currency is found for a risk-neutral person in three cases. It should be noted that another article [4, p. 9] addresses the problem, and it presents results different from those below. The case is that Slater solutions form a set the elements of which are in general different. Different elements of the same Slater set appear in [4] as well as in the present article. We will now proceed to the diversification problem.

The amount of money in a singular deposit diversified into two sub-deposits (in rubles and a foreign currency) accumulated by the end of the year may [18, p. 58–60] be represented as

$$\begin{gathered}\phi(x, y) = x(1 + r) + \frac{1 - x}{k}(1 + d)y = f_1(x) + f_2(x, y),\\ f_1(x) = x(1 + r), \quad f_2(x, y) = \frac{1 - x}{k}(1 + d)y,\end{gathered} \tag{8}$$

where r and d are the rates of interest for the sub-deposits in rubles and a foreign currency, accordingly; k and y are the exchange rates (to the ruble) in the beginning and the end of the year, accordingly; $x \in [0, 1]$ is a fraction defining the proportion, in which the singular deposit is divided into the sub-deposits. In accordance with (8), x is then the fraction of the sub-deposit in rubles, and the remaining part $1-x$ is converted into a foreign currency $\frac{1-x}{k}$ and allocated in a sub-deposit. In the end of the year it is converted back into rubles $\frac{1-x}{k}(1+d)y$ and the resulting amount of money is determined by the sum in (8). It is required to determine the fraction x, under the implementation of which the resulting amount of money $\phi(x,y)$ is the greatest possible one. It must be taken in account the future exchange rate y is normally unknown. We will, however, assume it may be defined in a corridor of possible values, precisely, $y \in [a,b]$, where the constants $b > a > 0$ are set or obtained at first. So, the mathematical model of the diversification problem in hand may be represented as an ordered tripod

$$\Gamma_2 = \langle X = [0,1], Y = [a,b], \{f_i(x,y)\}_{i=1,2} \rangle \,, \tag{9}$$

where function $f_i(x,y)$ $(i = 1,2)$ is defined in (8); here, $X = [0,1]$ is the set of strategies x of the decision maker; $Y = [a,b]$ is the set of uncertainties y; finally, $f_i(x,y)$ $(i = 1,2)$ is the function of usefulness (criterion), the value of which is called outcome. From the point of view operations research, Γ_2 is a two-criterion problem of making decisions under uncertainties. Presence of an uncertainty, as well as the intent to consider it is strongly tied to risks — "the possibility of deviation of any variables from the desired values" [7, p.15].

We will investigate three cases of the said problem:

first, $k\dfrac{1+r}{1+d} \leq a$;

second, $k\dfrac{1+r}{1+d} \geq b$;

third, $a \leq k\dfrac{1+r}{1+d} \leq b$.

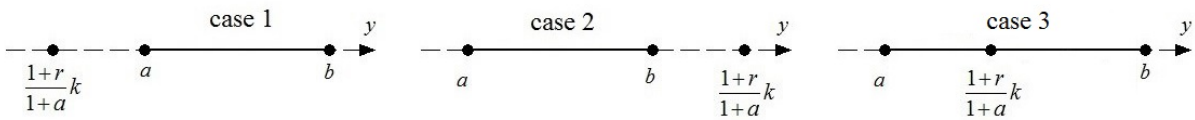

Fig. 1.

Cases 1 and 2.

Statement 2.

$$(x^S, f[x^S], R[x^S]) = (x^S, f_1[x^S], f_2[x^S], R_1[x^S], R_2[x^S]) =$$

$$= \begin{cases} \left(0, 0, \dfrac{1+d}{k}a, 1+r, 0\right), & k\dfrac{1+r}{1+d} \leq a, \\ \left(1, 1+r, 0, 0, \dfrac{1+d}{k}b\right), & k\dfrac{1+r}{1+d} \geq b. \end{cases} \tag{10}$$

Proof. **Part 1**.

Stage 1. Building the Savage–Nichans risk functions in accordance with (4):

$$R_1(x,y) = \left[\max_{z\in[0,1]} f_1(z)\right] - (1+r)x = (1-x)(1+r),$$

$$R_2(x,y) = \left[\max_{z\in[0,1]} f_1(z,y)\right] - (1-x)d\frac{1+r}{k}y = xy\frac{1+d}{k}.$$

Stage 2. Building guarantees of outcomes and risks. It is easy to see that

$$f_1[x] = \min_{y\in[a,b]} x(1+r) = x(1+r), \qquad f_2[x] = \min_{y\in[a,b]} (1-x)\frac{1+d}{k}y = (1-x)\frac{1+d}{k}a,$$

$$R_1[x] = \max_{y\in[a,b]} R_1(x,y) = (1-x)(1+r), \qquad R_2[x] = \max_{y\in[a,b]} R_2(x,y) = x\frac{1+d}{k}b.$$

Stage 3. Finding the optimal strategy in the four-criteria problem $< X = [0,1], \{f_i[x], -R_i[x]\}_{i=1,2}) >$. From this, we will also define criteria for the subsidiary problem $\Gamma = \left\langle X = [0,1], \{F_i[x]\}_{i=1,2} \right\rangle$:

$$F_1[x] = f_1[x] - R_1[x] = (2x-1)(1+r)$$
$$F_2[x] = f_2[x] - R_2[x] = \frac{1+d}{k}a - \frac{1+d}{k}(a+b)x$$

Part 2.

Stage 4. We now want to maximize the sum of criteria

$$\max_{x\in[0,1]} (F_1[x] + F_2[x]) = F_1[x^S] + F_2[x^S]$$

and build the Pareto maximal (and, therefore, Slater maximal) strategy x^S in accordance with

$$F[x^S] = \max_{x\in[0,1]} F[x],$$

where $F[x] = F_1[x] + F_2[x] = \dfrac{1+d}{k}\left(x(2\gamma - a - b) - (\gamma - a)\right)$, where $\gamma = \dfrac{1+r}{1+d}k$. The function to be maximized, $F[x]$, is linear with respect to x and defined in $[0,1]$ and, therefore, can only reach its maximal value only in either $x = 0$ or in $x = 1$; $F[0] = \dfrac{1+d}{k}(a-\gamma)$; $F[1] = \dfrac{1+d}{k}(\gamma - b)$.

Lemma 2. *The following implications are correct:*

$$a \geq \gamma \implies F[0] > F[1].$$

Proof. Indeed,

$$a \geq \gamma \implies \frac{a+a}{2} \geq \gamma \implies \frac{a+b}{2} > \gamma \implies a-\gamma > \gamma - b \implies$$
$$\implies F[0] = \frac{1+d}{k}(a-\gamma) > \frac{1+d}{k}(\gamma - b) = F[1].$$

□

Analogously,

Lemma 3. *The following implications are correct:*

$$\gamma \geq b \implies F[0] < F[1].$$

□

Stage 5. According to the lemmas above, maximum is attained under $x^S = 0$ when $a \geq \gamma$ and under $x^S = 1$ when $\gamma \geq b$. The corresponding guaranties can be obtained from here and Part 2: $f_1[0] = 0$, $f_2[0] = \frac{1+d}{k}a$, $R_1[0] = 1+r$, $R_2[0] = 0$, $f_1[1] = 1+r$, $f_2[1] = 0$, $R_1[1] = 0$, $R_2[1] = \frac{1+d}{k}b$.

We will now only note that $R_1[0] = 1+r$ is the risk of $f_1[0]$ equaling 0. $f[1] = 1+r$ is also less than the maximal outcome. That demonstrates the essence of the Savage–Nichans risk. Similar is the "meaning" of $R_2[1] = \frac{1+d}{k}b$. This result has been partially obtained by the authors in [19] using a different method than described above.

Case 3. Here we will use evidence for stage 3 of statement 2 proof in the first, two-criteria problem $\Gamma^\alpha = \langle [0,1], F_i[x]_{i=1,2} \rangle$, in particular, where

$$\begin{aligned} F_1[x] &= (2x-1)(1+r), \\ F_2[x] &= \frac{1+d}{k}a - (2x-1)(1+r); \end{aligned} \tag{11}$$

second, sufficient conditions (7) for existence of x^S from statement 1 which for problem (9) take the form

$$\min_{i=1,2} F_i[x^S] = \max_{x\in[0,1]} \min_{i=1,2} F_i[x].$$

Statement 3. *For $a < k\frac{1+r}{1+d} < b$, the SGOR is*

$$(x^S, f[x^S], R[x^S]) = (x^S; f_1[x^S], f_2[x^S]; R_1[x^S], R_2[x^S]) =$$

$$= \left(\frac{\gamma + a}{2\gamma + a + b}; \frac{\gamma + a}{2\gamma + a + b}(1+r), \frac{\gamma + b}{2\gamma + a + b} \frac{1+d}{k} a; \right. \tag{12}$$
$$\left. (1+r)\frac{\gamma + b}{2\gamma + a + b}, b\frac{1+d}{k}\frac{\gamma + a}{2\gamma + a + b} \right).$$

Proof. Plot two functions $F_1(x)$ and $F_2(x)$. These functions are linear with respect to x, defined on a compact $[0, 1]$, and their plots are shown on the Fig. 2.

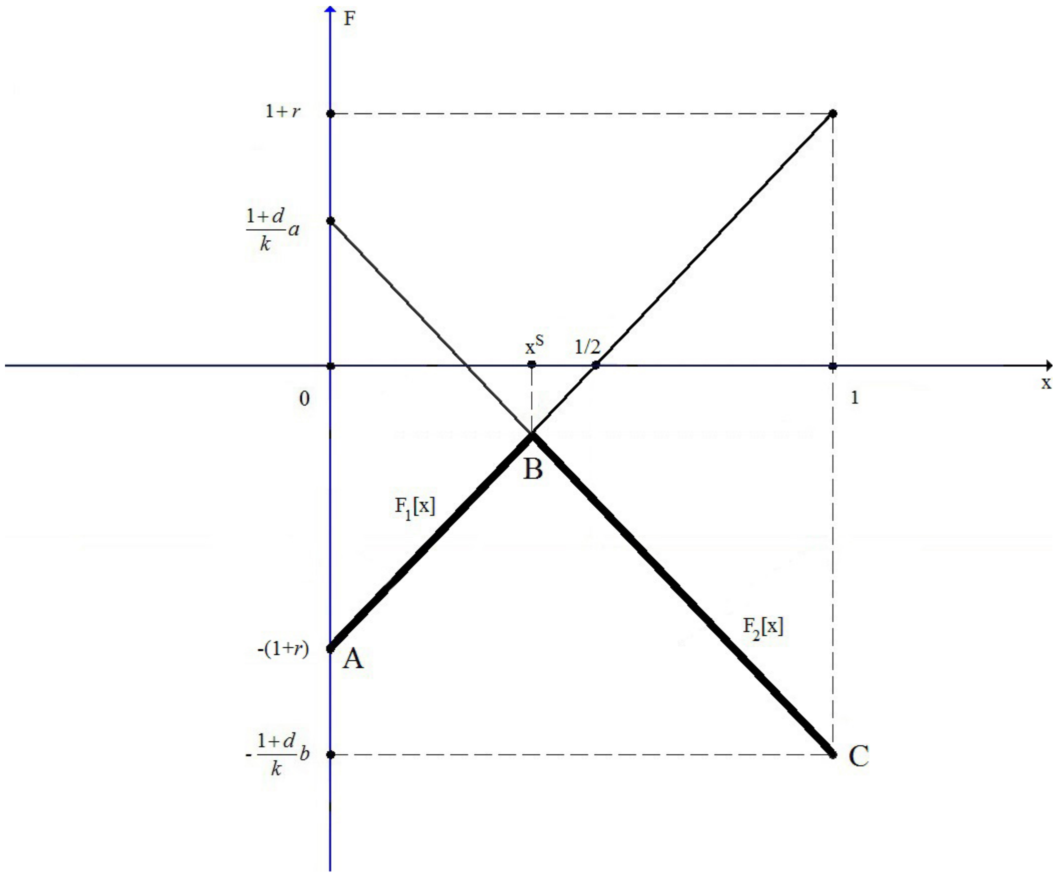

Fig. 2.

Function $\min_{i=1,2} \{F_1(x), F_2(x)\}$ is shown in bold on Fig. 2: the angle ABC. Point B is determined for $\max_{x\in[0,1]} \min_{i=1,2} \{F_1[x], F_2[x]\}$ by the equality

$$F_1(x^S) = F_2(x^S),$$

or given the functions $F_1[x]$ and $F_2[x]$,

$$x^S[2(1+r) + \frac{1+d}{k}(a+b)] = 1 + r + \frac{1+d}{k}a,$$

or, assuming $\gamma = \frac{1+r}{1+d}k$,

$$x^S[2\gamma + (a+b)] = \gamma + a,$$

from which follows

$$x^S = \frac{\gamma + a}{2\gamma + a + b},$$

and

$$1 - x^S = \frac{\gamma + b}{2\gamma + a + b},$$

Finally, using the formulas from step 2 of statement 2 proof we can find

$$f_1[x^S] = (1+r)\frac{\gamma + a}{2\gamma + a + b}, \qquad f_2[x^S] = a\frac{1+d}{k}\frac{\gamma + b}{2\gamma + a + b},$$
$$R_1[x^S] = (1+r)\frac{\gamma + b}{2\gamma + a + b}, \qquad R_2[x^S] = b\frac{1+d}{k}\frac{\gamma + b}{2\gamma + a + b}.$$

□

Acknowledgments

The authors thank the participants of the seminar “Risks in complex control systems” in the Computational Mathematics and Cybernetics Faculty of the Moscow State University for the discussion of the present work and their remarks.

The research has been funded by the Russian Foundation for Basic Research (grant 14-01-90408) and the National Academy of Sciences of Ukraine (grant 03-01-14).

References

1. Wald, A. (1939) Contribution to the theory of statistical estimation testing hypothesis. *Annuals Math. Statist.* 10. p. 299–326.

2. NICHANS, J. (1948) Zur Preisbildung bei ungewissen Erwartungen. *Schweizerische Zeitschrift für Volkswirtschaft and Statistic.* 84 (5). p. 433–456.

   NICHANS, J. (1948) On pricing of unknown expectations. *Swiss Journal for Popular Economics and Statistics.* 84 (5). p. 433–456.

3. SAVAGE, L. (1951) The theory of statistical division. *Journal of the American Statistical Association.* 46 (253). p. 55–67.

4. Черемных, Ю. Н. Микроэкономика. Продвинутый уровень. — М.: ИНФРА-М, 2008. — 843 с.

   CHEREMNYKH, Yu. (2008) *Microeconomics. Advanced level.* Moscow: INFRA-M.

5. Жуковский, В. И. Риски в конфликтных ситуациях. — М.: URSS, 2011. — 330 с.

   ZHUKOVSKIY, V. (2011) *Risks in conflict situations.* Moscow: URSS.

6. Сиразетдинов, Т. К., Сиразетдинов, Р. Т. Проблема риска и его моделирование // Проблемы человеческого риска. — 2007. — 1. — С. 31–43.

   SIRAZETDINOV, T. & SIRAZETDINOV, R. (2007) Problem of risc and and its modeling. *The problem of human risk.* 1. p. 31–43.

7. Шахов, В. В. Введение в страхование. Экономический аспект. — М.: Финансы и статистика, 2001. — 286 с.

   SHAKHOV, V. (2001) *Introduction into insurance. Economical aspect.* Moscow: Finansy i statistika.

8. Цветкова, Е. В., Арлюкова, Н. О. Риск в экономической деятельности. — СПб.: ИВЭСЭП, 2002. — 64 с.

   TSVETKOVA, E. and ARLUKOVA, N. (2002) *Risk in economic activity.* SPb: IEECEL.

9. MARKOWITZ, H. (1952) Portfolio Selection. *Journal of Finance.* Vol. 7 (№ 1, March). p. 55–67.

10. Фишер, С., Дорнбушб, Р., Шмалензи, Р. Экономика. — М.: Дело, 1998. — 829 с.

    FISHER, S., DORNBUSH, R and SHMALENZI, R. (1998) *Economy.* Moscow: Delo.

11. Жуковский, В. И., Кудрявцев, К. Н. Уравновешивание конфликтов при неопределенности. I. Аналог седловой точки // Математические основы теории игр и приложения. — 2013. — 5. 1. — С. 27–44.

    ZHUKOVSKIY, V. and KUDRYAVTSEV, K. (2013) Solving conflicts under uncertainty. I. Analog of saddle point. *Mathematical foundation of game theory and applications.* 5 (1). p. 27–44.

12. Жуковский, В. И., Кудрявцев, К. Н. Уравновешивание конфликтов при неопределенности. II. Аналог максимина // Математические основы теории игр и приложения. — 2013. — 5. 2. — С. 3–45.

    ZHUKOVSKIY, V. and KUDRYAVTSEV, K. (2013) Solving conflicts under uncertainty. II. Analog of maximin. *Mathematical foundation of game theory and applications.* 5 (2). p. 3–45.

13. ZHUKOVSKIY, V. and SALUKVADZE, M. (1994) *The Vector — Valued Maximin.* New York: Academic Press.

14. Морозов, В. В., Сухарев, А. Г., Федоров, В. В. Исследование операций в задачах и упражнениях. — М.: Высшая школа, 1986. — 286 с.

    MOROZOV, V., SUKHAREV, A. and FEDOROV, V. (1968) *Operational research in problems and exercises.* Moscow: Higher school.

15. Подиновский, В., В., Ногин, В. Д. Парето–оптимальное решение многокритериальных задач / В. В. Подиновский, В. Д. Ногин. — М.: Физматлит, 2007. — 256 с.

    PODINOVSKIY, V. and NOGIN, V. (2007) *Pareto optimal solution of multicriteria problems.* Moscow: Fizmatlit.

16. Эрроу, К. Дж., Гурвиц, Л., Удзава, Х. Исследования по линейному и нелинейному программированию. — М.: Изд-во инстранной литературы, 1962. — 334 с.

    ARROW, K., HURWICS, L. and UZAWA H. (1974) *Research in linear and nonlinear programming.* Moscow: Publishing house of foreign literature.

17. ZHUKOVSKIY, V., MOLOSTVOV, V. and TOPCHISHVILI, A. (2014) Problem of multicurrency deposit diversification – three possible approaches to risk accounting. *International Journal of Operations and Quantitative Management.* 20 (1). p. 1–15.

18. Капитоненко, В. В. Финансовая математика и ее приложения. — М.: ПРИОР, 2000. — 140 с.

    KAPITONENKO, V. (2000) *Fiscal mathematics and its applications.* Moscow: PRIOR.

19. Жуковский, В. И., Кириченко, М. М. Риски и исходы в многокритериальной задаче при неопределенности // Управление риском. — 2016. — Т. 78. — С. 17–28.

    ZHUKOVSKIY, V. & KIRICHENKO, M. (2016) Risk and outcomes in multicriteria problem under uncertainty. *Risk management.* Vol. 78 (2). p. 17–28.

УДК 519.95
MSC2010: 68Q32

УСТОЙЧИВОСТЬ АЛГОРИТМОВ ОБУЧЕНИЯ КЛАССИФИКАЦИИ, ОСНОВАННЫХ НА МОДИФИЦИРОВАННОЙ МОДЕЛИ ВЫЧИСЛЕНИИ ОЦЕНОК

© А. С. Анафиев, В. Ф. Блыщик, В. И. Донской

Крымский федеральный университет им. В. И. Вернадского
факультет математики и информатики
пр-т Академика Вернадского, 4, г. Симферополь, 295007, Россия
e-mail: *aydera@mail.ru, donskoy@tnu.crimea.ua*

Stability of Learning Classification Algorithms Based on the Modified Estimates Calculation Model.

Anafiyev A. S., Blyschik V. F., Donskoy V. I.

Abstract. In this paper, we obtain the following theoretical result: there exists a stable algorithm $\mathscr{A}$ of the modified model ABO^* training, guaranteeing its learnability in the form of universal empirical generalization directly by use learning sample by the way minimizing the empirical risk. To obtain this result the LOO stability of the algorithm $\mathscr{A}$ was proved. Algorithm $\mathscr{A}$ described in details in this article is a learning procedure with adaptation. It requires the adjustment of only the weights of objects of training sample. The remaining parameters of the model remain fixed. This is sufficient to achieve the desired result. Proposed modification of the model ABO is minimal: it excludes only the case where "the point is voting for itself". It is easy to show that in the case when the modified model ABO^* is learned by only the choice the shortest elementary logical separators (in particular — a dead-end tests), a universal empirical generalization will also take place.

1. Введение. Постановка задачи

В теории машинного обучения центральным вопросом является *проблема обучаемости*, под которой понимают способность алгоритмов обучения к эмпирическому обобщению — свойству, заключающемуся в их способности по конечному множеству частных случаев (объектов, примеров, прецедентов) построить обобщение в виде решающего правила, способного гарантированно обеспечить высокую вероятность правильного принятия решения при предъявлении любого объекта, не участвовавшего в обучении.

В рамках этой статьи в качестве частного случая, достаточного для получения основного теоретического результата, взята *задача обучения классификации объектов для случая двух классов*.

Целью работы является доказательство возможности построения устойчивого алгоритма обучения классификации в рамках модифицированной модели вычисления оценок (ABO) на основе метода минимизации эмпирического риска по *одной заданной обучающей выборке*. Модификация модели, на первый взгляд, может показаться незначительной, и она, действительно, минимальна: изменяется лишь область суммирования в оценках, исключающая вклад в суммы ровно одной точки.

Модель ABO была создана академиком Ю. И. Журавлёвым [11, 13, 12] и развивала идеи предложенных им алгоритмов распознавания, основанных на построении тупиковых тестов таблиц обучения как совокупности оптимальных (по минимуму длины) элементарных классификаторов [6]. В дальнейшем теория моделей ABO получила широкое развитие в работах академиков В. Л. Матросова [18, 19] и К. В. Рудакова [2, 21], профессоров В. В. Рязанова [22], А. Г. Дьяконова [8, 9, 10], а также многих других исследователей [16, 20, 23, 17, 15].

Первоначальные модели ABO предполагали тщательный анализ таблиц прецедентов и позволяли автоматически получать функции принадлежности объектов классам. Заметим, что в этом смысле был получен важный результат, на основе которого стало возможным строить эмпирические нечёткие модели принятия решений, поскольку именно получение функций принадлежности, а не построение операций над ними, являлось "узким местом" теории нечётких множеств [14].

Являясь по сути эталонными, в случае точной начальной прецедентной информации модели ABO на уровне интуитивного представления должны были давать все более высокую точность с ростом числа эталонов, поскольку использовали не полное совпадение с ними, а аппроксимирующую функцию принадлежности, вычисляемую на основе специально введенного расстояния между объектами. Однако при непосредственном применении метода минимизации эмпирического риска по эталонной выборке статистическая теория обучения Вапника–Червоненкиса давала отрицательный результат об обучаемости, поскольку в указанной постановке ёмкость (VC размерность) модели ABO оказывалась бесконечной.

Более поздние теоретические результаты, показавшие, что обучаемость может быть обеспечена не только за счёт конечной ёмкости применяемого семейства решающих правил, но и за счёт устойчивости алгоритмов обучения [24, 25, 26, 27], послужили поводом для дополнительного изучения вопроса об обучаемости модели ABO методом минимизации эмпирического риска непосредственно по одной обучающей выборке.

В статье используются следующие основные обозначения:

$(x_1, \ldots, x_n) = \tilde{x} \in X^n \subset \mathbb{R}^n$ — точка или допустимый объект, описанный n вещественными переменными–признаками;

K_0 и K_1 — два класса (множества) объектов, $K_0 \cup K_1 = X^n$; $\alpha \in \{0; 1\}$ — номер класса K_α;

$(\tilde{x}, \alpha)$ — прецедент или пример, являющийся парой, содержащей некоторый объект $\tilde{x}$ с заведомо и точно известным номером класса α, которому этот объект принадлежит ($\tilde{x} \in K_\alpha$);

$X_l = \{(\tilde{x}_j, \alpha_j)_{j=1}^l\}$ — обучающая выборка, состоящая из l прецедентов, представляющая собой набор представителей двух классов: $X_l = T_0 \cup T_1;\ T_0 \cap T_1 = \emptyset;\ T_0 \subset K_0;\ T_1 \subset K_1$;

$\mathscr{X}_l$ — множество всевозможных обучающих выборок, содержащих l прецедентов (имеющих длину l);

$\mathfrak{F} = \{\varphi : X^n \to \{0;1\}\}$ — произвольное семейство решающих правил (классификаторов);

$A : \mathscr{X}_l \to \mathfrak{F}$ — произвольный алгоритм (или метод) обучения;

$A(X_l) = h \in \mathfrak{F}$ — алгоритм классификации, обученный по данной выборке X_l;

$A(X_l, \tilde{\theta})$ — алгоритм классификации, обученный по данной выборке X_l в случае, если $\mathfrak{F}$ — параметрическое семейство; $\tilde{\theta}$ — набор параметров;

$A(X_l)(\tilde{x}) \in \{0;1\}$ — результат классификации объекта $\tilde{x}$ алгоритмом $A(X_l)$, обученным по выборке X_l;

$\nu(A(X_l)) = \frac{1}{l}|\{\tilde{x}_j : A(X_l)(\tilde{x}_j) \neq \alpha_j\}_{j=1}^l|$ — частота ошибок при классификации примеров выборки X_l алгоритмом $A(X_l)$, обученным по этой же выборке X_l (в общем случае не все примеры выборки могут классифицироваться правильно);

$\mathrm{P}(A(X_l))$ — вероятность ошибки (неизвестная) алгоритма $A(X_l)$, обученного по произвольной выборке длины l; $\mathrm{P}(A(X_l)) = \mathbf{E}_P\Big[\big|A(X_l)(\tilde{x}) \neq \alpha(\tilde{x})\big|\Big]$, где $\alpha(\tilde{x})$ — истинный, но неизвестный номер класса, которому принадлежит точка $\tilde{x}$;

$Err_{\mathrm{P}}(A(X_l)) = \mathrm{P}(A(X_l))$ — эквивалентное обозначение вероятности ошибки алгоритма $A(X_l)$;

$\mathbf{E}_P$ — математическое ожидание по мере P;

P — вероятностное распределение (неизвестное) на множестве допустимых примеров $X^n \times \{0;1\}$;

$\mathscr{P}^l$ — вероятностное распределение (неизвестное) на множестве $\mathscr{X}_l = \big(X^n \times \{0;1\}\big)^l$ всевозможных обучающих выборок длины l;

λ — функция потерь: $\lambda(h, \tilde{x}) = \begin{cases} 0, & \text{если } h(\tilde{x}) = \alpha(\tilde{x}); \\ 1, & \text{если } h(\tilde{x}) \neq \alpha(\tilde{x}) \text{ или } h(\tilde{x}) \text{ не определено.} \end{cases}$

Следующие два параграфа имеют обзорный характер и предназначены для разъяснения ряда используемых определений и теорем теории машинного обучения, которые применялись в процессе получения основного результата работы.

2. VC размерность и обучаемость

В теории машинного обучения одним из важнейших понятий является *VC-размерность или емкость семейств отображений*, из которых извлекаются решающие правила [3]. С этим понятием тесно связано представление об обучаемости алгоритмов в форме равномерной по классу используемых классификаторов сходимости эмпирических частот ошибок обученных классификаторов к их вероятностям. И поскольку данная статья посвящена обучаемости модифицированной модели ABO, ниже со ссылками на первоисточник приводятся основные положения статистической теории обучения Вапника–Червоненкиса.

Применение произвольного классификатора $\varphi \in \mathfrak{F}$ к l точкам из выборки X_l порождает l двоичных значений — бинарную строку

$$\tilde{y} = (y_1, \ldots, y_j, \ldots, y_l) \ : y_j = \varphi(\tilde{x}_j) \in \{0,1\},\ j = 1, \ldots, l.$$

Будем называть строку $\tilde{y}$ *разбиением выборки* X_l *на два класса* в соответствии со значениями 0 и 1 классификатора φ и использовать обозначение $\tilde{y}_\varphi = \tilde{\varphi}(\tilde{X}_l)$.

Применение одного и того же алгоритма к различным выборкам и применение различных алгоритмов к одной и той же выборке даёт, вообще говоря, различные разбиения. Алгоритмы–классификаторы семейства $\mathfrak{F}$, порождающие одинаковые разбиения любых допустимых выборок, будем называть *подклассом эквивалентных алгоритмов* семейства $\mathfrak{F}$.

Определение 1. [3] Пусть $\Delta^{\mathfrak{F}}(\tilde{x}_1, \ldots, \tilde{x}_l)$ — число различных классификаций выборки X_l на два класса, получаемое при использовании всех классификаторов семейства $\mathfrak{F}$. Функцией роста семейства $\mathfrak{F}$ называется

$$m^{\mathfrak{F}}(l) = \max_{\tilde{x}_1, \ldots, \tilde{x}_l} \Delta^{\mathfrak{F}}(\tilde{x}_1, \ldots, \tilde{x}_l),$$

где максимум берётся по всем возможным последовательностям из l точек (допустимых объектов).

Определение 2. Энтропией семейства $\mathfrak{F}$ на обучающих выборках длины l называется величина

$$H^{\mathfrak{F}}(l) = \mathbf{E}_{\mathscr{P}^l}\big(\ln \Delta^{\mathfrak{F}}(\tilde{x}_1, \ldots, \tilde{x}_l)\big).$$

В [3] доказано, что функция роста $m^{\mathfrak{F}}(l)$ либо тождественно равна 2^l, либо, если это не так, мажорируется функцией $\sum_{i=0}^{h} C_l^i \leq 1.5\frac{l^h}{h!}$. Если это неравенство имеет место, то минимальное удовлетворяющее ему число h называют ёмкостью семейства классифицирующих функций $\mathfrak{F}$.

Определение 3. [3] VC-размерностью или ёмкостью семейства классифицирующих функций $\mathfrak{F} = \{\varphi : X^n \to \{0,1\}\}$, обозначаемой $VCD(\mathfrak{F})$, называется наибольшее значение l^* такое, что найдется выборка X_{l^*}, которая может быть разбита всеми 2^{l^*} способами алгоритмами семейства $\mathfrak{F}$:

$$\exists X_{l^*} \ : \ \forall \tilde{y} \in \{0,1\}^{l^*} \ \ \exists \varphi \in \mathfrak{F} \ \big(\tilde{y} = \tilde{\varphi}(\tilde{X}_{l^*})\big),$$

но никакая выборка длины большей, чем l^*, разбита функциями этого семейства всеми способами быть не может.

Если же при любом l найдется выборка, разбиваемая всеми 2^l способами, то VC-размерность семейства $\mathfrak{F}$ полагается неограниченной (равной ∞).

Теорема 1. [3] *Вероятность того, что хотя бы для одного классификатора φ, принадлежащего семейству $\mathfrak{F}$ и выбранного методом минимизации эмпирического риска, частота ошибки на обучающей выборке длины l отклонится от её вероятности более, чем на любое* *положительное $\varepsilon > 0$, удовлетворяет неравенствам*

$$\mathscr{P}^l\left(\sup_{\varphi\in\ \mathfrak{F}}\left|\mathrm{P}(\varphi(X_l)) - \nu(\varphi(X_l))\right| > \varepsilon\right) < 6m^{\mathfrak{F}}(2l)e^{-\frac{\varepsilon^2 l}{4}};$$

$$\mathscr{P}^l\left(\sup_{\varphi\in\ \mathfrak{F}}\left|\mathrm{P}(\varphi(X_l)) - \nu(\varphi(X_l))\right| > \varepsilon\right) < 9\frac{(2l)^{VCD(\mathfrak{F})}}{VCD(\mathfrak{F})\ !}(2l)e^{-\frac{\varepsilon^2 l}{4}}.$$

Следствие 1. *Для того, чтобы частота ошибки на обучающей выборке любого решающего правила φ, принадлежащего семейству $\mathfrak{F}$ и выбранного методом минимизации эмпирического риска равномерно сходилась (по вероятности) при $l \to \infty$ к вероятности ошибки этого правила φ, достаточно, чтобы ёмкость $VCD(\mathfrak{F})$ семейства $\mathfrak{F}$ была конечной.*

Иначе говоря, конечность $VCD(\mathfrak{F})$ семейства $\mathfrak{F}$ является *достаточным условием обучаемости*, понимаемой как гарантированная возможность найти путём минимизации эмпирической ошибки решающее правило $\varphi \in \mathfrak{F}$ такое, что его истинная вероятность ошибки гарантированно не будет превышать эмпирическую ошибку на сколь угодно малое положительное $\varepsilon > 0$ со сколь угодно высокой надёжностью при неограниченном росте длины обучающей выборки.

Необходимым условием обучаемости (в том же смысле, как это оговаривается в достаточном условии) является условие [3]

$$\lim_{l\to\infty}\frac{H^{\mathfrak{F}}(l)}{l} = 0.$$

Если $VCD(\mathfrak{F}) = \infty$, то для любого l достоверно существует хотя бы одна выборка X_l, которую можно расклассифицировать всеми 2^l способами. И если в пространстве всех выборок такие выборки имеют ненулевую меру, то можно показать, что найдётся константа c такая, что

$$\lim_{l\to\infty}\frac{H^{\mathfrak{F}}(l)}{l} \to c > 0.$$

Поэтому необходимое условие обучаемости по Вапнику–Червоненкису, требующее сколь угодно малой энтропии на символ для класса $\mathfrak{F}$, является близким к требованию конечности $VCD(\mathfrak{F})$.

Определение 4. Семейство классификаторов $\mathfrak{F}$ называется PAC (*Probably Approximately Correct*) обучаемым (при использовании семейства гипотез–классификаторов $\mathscr{H}$), если существует алгоритм обучения A, который на основе прецедентного описания любого классификатора $\varphi \in \mathfrak{F}$ в виде обучающей выборки длины l, при любых вероятностных распределениях $\mathscr{P}^l$ (обучающих выборок) и P(допустимых точек-объектов признакового пространства с указанной принадлежностью классам) определяет гипотезу $h \in \mathscr{H}$ такую, что для любых $\varepsilon, \delta :\ 0 < \varepsilon,\, \delta < \frac{1}{2}$

$$\mathscr{P}^l[Err_{\mathrm{P}}(h) < \varepsilon] > 1 - \delta, \tag{1}$$

и при этом существует функция $l = l(\varepsilon, \delta)$, которая обеспечивает выполнение неравенства (1), где $Err_{\mathrm{P}}(h) = \mathrm{P}[h(\tilde{x}) \neq \varphi(\tilde{x})]$.

Вариант модели PAC обучаемости, когда целевой неизвестный классификатор $\varphi \in \mathfrak{F}$ заведомо содержится в семействе H, используемом для обучения, называется реализуемой PAC моделью (или правильной PAC обучаемостью) [25].

Теорема 2. [24, 27] *Семейство классификаторов $\mathfrak{F}$ является PAC обучаемым тогда и только тогда, когда $VCD(\mathfrak{F}) < \infty$.*

Несмотря на то, что известны различные подходы к формулировке обучаемости, во многих моделях требование конечности VC размерности семейства, в котором отыскивается классификатор или которое содержит искомый классификатор, является определяющим.

3. Устойчивость алгоритмов обучения и обучаемость

В этом параграфе обзорно даётся одна из формализаций устойчивости обучающих алгоритмов. Приводятся известные результаты [26], необходимые для решения поставленной задачи — исследования устойчивости и обучаемости модифицированной модели ABO.

Определение 5. Говорят, что *имеет место универсальное эмпирическое обобщение*, если для любого выбранного алгоритмом обучения A классификатора (гипотезы) частота ошибки этого классификатора на обучающей выборке сходится по вероятности к её математическому ожиданию при неограниченном росте длины обучающей выборки независимо от вероятностного распределения, т. е.

$$\forall \varepsilon > 0 \quad \mathscr{P}^l \Big\{ \big| \mathrm{P}\big(A(X_l)\big) - \nu\big(A(X_l)\big) \big| > \varepsilon \Big\} \to 0 \quad \text{при} \quad l \to \infty \tag{2}$$

для любой гипотезы $A(X_l)$ и любых вероятностных мер $\mathscr{P}^l$ и P.

Уточним, что в формуле для эмпирической частоты ошибки

$$\nu\big(A(X_l)\big) = \frac{1}{l} \sum_{j=1}^{l} \lambda\big(A(X_l)(\tilde{x}_j), \alpha(\tilde{x}_j)\big)$$

$A(X_l)(\tilde{x}_j)$ — результат классификации точки $\tilde{x}_j$ из примера $(\tilde{x}_j, \alpha)$ обучающей выборки X_l алгоритмом $A(X_l)$, а $\alpha(\tilde{x}_j) = \alpha$ — класс, к которому относится точка $\tilde{x}_j$ согласно обучающему примеру $(\tilde{x}_j, \alpha)$.

Вероятность ошибки обученного алгоритма $A(X_l)$, обозначенная в (2) как $\mathrm{P}\big(A(X_l)\big)$, есть вероятностная мера события-ошибки, состоящей в несовпадении результата классификации произвольной точки $\tilde{x}$ алгоритмом $A(X_l)$ с истинной, но неизвестной классификацией этой точки. Введенная выше функция потерь λ является характеристической функцией ошибки, поэтому

$$\mathrm{P}\big(A(X_l)\big) = Err\big(A(X_l)\big) =$$
$$= \int\limits_{X^n \times \{0;1\}} \lambda\big(A(X_l)(\tilde{x}), \alpha(\tilde{x})\big) dP(\tilde{x}, \alpha) = \mathbf{E}_P \Big(\lambda\big(A(X_l)(\tilde{x}), \alpha(\tilde{x})\big) \Big),$$

где **E** — символ математического ожидания, и в последнем выражении используется интеграл Лебега.

В определении 5 семейство решающих правил, в котором отыскивается классификатор $A(X_l)$, не содержится в явном виде. *Может оказаться, что алгоритм обучения* (метод обучения по К. В. Воронцову [5]) *обладает свойством "сужения" семейства, в котором в принципе могут отыскиваться гипотезы*:

$$A : \mathscr{X}_l \to \mathfrak{F}_A \subset \mathfrak{F}, \quad \mathfrak{F}_A \neq \mathfrak{F}. \tag{3}$$

Поэтому VC размерность семейства $\mathfrak{F}$, вообще говоря, может быть любой, и при этом возможно универсальное эмпирическое обобщение. В то же время сходимость (2) обеспечивается для любой гипотезы $A(X_l)$ и тогда

$$\forall (A(X_l)\Big(\big|\mathrm{P}\big(A(X_l)\big) - \nu\big(A(X_l)\big)\big| > \varepsilon\Big) \Rightarrow \left(\sup_{\{A(X_l)\in\ \mathfrak{F}_A\}}\big|\mathrm{P}\big(A(X_l)\big) - \nu\big(A(X_l)\big)\big| > \varepsilon\right),$$

но при этом

$$\left(\sup_{\{A(X_l)\in\ \mathfrak{F}\}}\big|\mathrm{P}\big(A(X_l)\big) - \nu\big(A(X_l)\big)\big| > \varepsilon\right) \Rightarrow \Big(\forall A(X_l)\big(\big|\mathrm{P}\big(A(X_l)\big) - \nu\big(A(X_l)\big)\big| > \varepsilon\big)\Big).$$

В определении Вапника–Червоненкиса равномерной сходимости эмпирической частоты к вероятности выбранного классификатора *метод обучения вообще не фигурирует*, но при выполнении соотношений (3) справедлива принадлежность $A(X_l) = \varphi \in \mathfrak{F}$, поэтому при выполнении условия (3) равномерная сходимость по классу $\mathfrak{F}$ влечёт универсальное эмпирическое обобщение. Следовательно, конечность $VCD(\mathfrak{F})$ при выполнении (3) является достаточным условием наличия универсального эмпирического обобщения; но, как будет показано ниже, не является необходимым.

Обозначим X_l^j обучающую выборку X_l, из которой *удалён ровно один пример* $(\tilde{x}_j, \alpha_j)$.

Определение 6. [26] Алгоритм обучения A называется CV_{Loo} устойчивым (*Cross-Validation Leave-one out*) независимо от распределения, если для любой вероятностной меры, для любой длины выборки $l \geq l_0$ найдутся такие положительные $\varepsilon(l), \delta(l) < 1$, что при $l > l_0,\ l \to \infty,\ \ \varepsilon(l) \to 0,\ \ \delta(l) \to 0$

$$\forall j \in \{1, \ldots, l\} \ \ \mathscr{P}^l\big(|\lambda(A(X_l^j), \tilde{x}_j) - \lambda(A(X_l), \tilde{x}_j)| \leq \varepsilon(l)\big) \geq 1 - \delta(l),$$

где λ — функция потерь.

Определение 7. [26] Алгоритм обучения A называется $ELoo_{err}$ устойчивым независимо от распределения, если для любой вероятностной меры при любом значении $l > l_0$ найдутся такие положительные $\varepsilon(l), \delta(l) < 1$, что

$$\forall j \in \{1, \ldots, l\} \ \mathscr{P}^l\Big(|Err(A(X_l)) - \frac{1}{l}\sum_{j=1}^{l}\lambda(A(X_l^j), \tilde{x}_j)| \leq \varepsilon(l)\Big) \geq 1 - \delta(l),$$

где $\varepsilon(l) \to 0$, $\delta(l) \to 0$ при $l \to \infty$; $Err(A(X_l))$ — определённая выше вероятность ошибки классификатора $A(X_l)$, обученного по данной выборке X_l; λ — определённая выше функция потерь.

Определение 8. [26] Алгоритм обучения A называется LOO устойчивым, если он одновременно CV_{Loo} устойчивый и $ELoo_{err}$ устойчивый.

Определение 9. Алгоритм обучения называется *симметричным*, если результат его применения к любой допустимой обучающей выборке не изменяется при любой перестановке входящих в эту выборку примеров.

Следующая теорема [26] особо важна для изложения главного результата статьи и поэтому приводится с подробным доказательством.

Теорема 3. *LOO устойчивость симметричного алгоритма обучения классификации с ограниченной функцией потерь является достаточным условием для обеспечения универсального эмпирического обобщения.*

Доказательство. Для упрощения формул обозначим $h = A(X_l)$ и

$$Err_l\big(A(X_l)\big) = Err_l(h) = \frac{1}{l}\sum_{j=1}^{l} \lambda\big(A(X_l), \tilde{x}_j\big)$$

эмпирическую ошибку обученного по выборке X_l алгоритма $A(X_l) = h$ в среднем по этой же выборке.

Оценим математическое ожидание (вероятностную меру) квадрата отклонения вероятности ошибки решающего правила (гипотезы) $h = A(X_l)$, полученного в результате обучения, от эмпирической ошибки этой гипотезы. Распределение $\mathscr{P}^l$, и семейство H, которому принадлежит гипотеза h, полагаются произвольными.

$$\begin{aligned}
&\mathscr{P}^l\Big(Err\big(A(X_l)\big) - \frac{1}{l}\sum_{j=1}^{l} \lambda\big(A(X_l), \tilde{x}_j\big)\Big)^2 \\
&= \mathbf{E}_l\Big(Err(h) - Err_l\big(A(X_l)\big)\Big)^2 \\
&= \mathbf{E}_l\Big(Err(h) - \frac{1}{l}\sum_{j=1}^{l} \lambda\big(A(X_l^j), \tilde{x}_j\big) + \frac{1}{l}\sum_{j=1}^{l} \lambda\big(A(X_l^j), \tilde{x}_j\big) - Err_l\big(A(X_l)\big)\Big)^2 \\
&\le 2\mathbf{E}_l\Big(Err(h) - \frac{1}{l}\sum_{j=1}^{l} \lambda\big(A(X_l^j), \tilde{x}_j\big)\Big)^2 \\
&+ 2\mathbf{E}_l\Big(\frac{1}{l}\sum_{j=1}^{l} \lambda\big(A(X_l^j), \tilde{x}_j\big) - Err_l(h)\Big)^2. \qquad (4)
\end{aligned}$$

Последнее неравенство следует из того, что $(a+b)^2 \le 2a^2 + 2b^2$.

Оценим второе слагаемое (4).

$$2\mathbf{E}_l\Big(\frac{1}{l}\sum_{j=1}^{l}\lambda\big(A(X_l^j),\tilde{x}_j\big) - Err_l(h)\Big)^2$$

$$= 2\mathbf{E}_l\Big(\frac{1}{l}\sum_{j=1}^{l}\lambda\big(A(X_l^j),\tilde{x}_j\big) - \frac{1}{l}\sum_{j=1}^{l}\lambda\big(A(X_l),\tilde{x}_j\big)\Big)^2$$

$$= 2\mathbf{E}_l\Big(\frac{1}{l}\sum_{j=1}^{l}\lambda\big(A(X_l),\tilde{x}_j\big) - \frac{1}{l}\sum_{j=1}^{l}\lambda\big(A(X_l^j),\tilde{x}_j\big)\Big)^2$$

$$= 2\mathbf{E}_l\frac{1}{l^2}\Bigg|\sum_{j=1}^{l}\Big(\lambda\big(A(X_l),\tilde{x}_j\big) - \lambda\big(A(X_l^j),\tilde{x}_j\big)\Big)\Bigg|^2 \tag{5}$$

$$\leq 2M\mathbf{E}_l\frac{1}{l}\Bigg|\sum_{j=1}^{l}\Big(\lambda\big(A(X_l),\tilde{x}_j\big) - \lambda\big(A(X_l^j),\tilde{x}_j\big)\Big)\Bigg|$$

(последний шаг доказательства основан на использовании ограниченности функции потерь, в силу чего $\Big|\sum_{j=1}^{l}\Big(\lambda\big(A(X_l),\tilde{x}_j\big) - \lambda\big(A(X_l^j),\tilde{x}_j\big)\Big)\Big| \leq M\cdot l$, где M — константа; в нашем случае — при бинарной функции потерь λ — имеем $M = 1$, и из квадрата модуля в (5) заменяем один модуль на l)

$$\leq 2\mathbf{E}_l\frac{1}{l}\sum_{j=1}^{l}\big|\lambda\big(A(X_l),\tilde{x}_j\big) - \lambda\big(A(X_l^j),\tilde{x}_j\big)\big|$$

$$= 2\frac{1}{l}\sum_{j=1}^{l}\mathbf{E}_l\big|\lambda\big(A(X_l),\tilde{x}_j\big) - \lambda\big(A(X_l^j),\tilde{x}_j\big)\big|$$

(далее учитываем, что A — симметричный алгоритм; $\mathbf{E}_l|\cdot|$ — математическое ожидание по вероятностному распределению $\mathscr{P}^l$ на множестве обучающих выборок не зависит от j)

$$= 2\mathbf{E}_l\big|\lambda\big(A(X_l),\tilde{x}_j\big) - \lambda\big(A(X_l^j),\tilde{x}_j\big)\big|$$

для любого примера $\tilde{x}_j$ (для любого j) из произвольной обучающей выборки X_l.

Окончательно получаем неравенство

$$\mathbf{E}_l\Big(Err(h) - Err_l\big(A(X_l)\big)\Big)^2$$

$$\leq 2\mathbf{E}_l\Big(Err(A(X_l)) - \frac{1}{l}\sum_{j=1}^{l}\lambda\big(A(X_l^j),\tilde{x}_j\big)\Big)^2$$

$$+2\mathbf{E}_l\big|\lambda\big(A(X_l),\tilde{x}_j\big) - \lambda\big(A(X_l^j),\tilde{x}_j\big)\big|,$$

в правой части которого содержатся два слагаемых, справедливых для любого $j \in \{1,\ldots,l\}$. Первое слагаемое соответствует определению $ELoo_{err}$ устойчивости, а второе — CV_{Loo} устойчивости. Если оба эти слагаемые при $l \to \infty$ одновременно стремятся к нулю, то, согласно определению, имеет место LOO устойчивость, что влечёт универсальное эмпирическое обобщение, поскольку сумма указанных слагаемых является верхней оценкой вероятности квадрата отклонения вероятности ошибки полученной в результате обучения гипотезы, от её эмпирической ошибки. □

4. Определение модели ABO и её необучаемость непосредственно по заданной выборке (без дополнительной "контрольной")

Будем полагать, что обучающая выборка состоит из двух частей — представителей (примеров) двух непересекающихся классов объектов K_0 и K_1 соответствующих выборочным значениям 0 и 1 классифицирующей функции:

$$X_l = \{(\tilde{x}_j, \alpha_j)_{j=1}^l\} = T_0 \cup T_1;\ T_0 \cap T_1 = \emptyset;$$

$$T_0 = \{(\tilde{x}, \alpha) : \alpha = 0\};\ T_1 = \{(\tilde{x}, \alpha) : \alpha = 1\};$$

$$|X_l| = l;\ |T_0| = k_0;\ |T_1| = k_1.$$

Потребуем, чтобы в обучающей выборке X_l не содержалось одинаковых точек (что легко обеспечивается исключением повторов и противоречий).

Метод (алгоритм) вычисления оценок (ABO), предназначенный для построения классификатора по заданной обучающей выборке, определяется следующим образом.

1. Точке $\tilde{x}$ каждого примера $(\tilde{x}, \alpha)$ обучающей выборки ставится в соответствие неотрицательное число $\omega(\tilde{x})$ — "*вес*" этого примера (*эталона* $\tilde{x}$).

2. Задаётся система множеств $\tilde{\Omega}$, называемых *опорными*, которые являются некоторым образом отобранными подмножествами множества $\{1, ..., n\}$ номеров переменных. Каждому опорному множеству $\Omega \in \tilde{\Omega}$ ставится в соответствие неотрицательное число $W(\Omega)$ — "*вес*" *опорного множества.*

3. Вводится расстояние между координатами x_i и y_i точек $\tilde{x}$ и $\tilde{y}$ по формуле $\rho(x_i, y_i) = |x_i - y_i|,\ i = \overline{1, n}$.

4. Определяется функция близости по опорному множеству:

$$B_\Omega(\tilde{x}, \tilde{y}) = \begin{cases} 1, & \text{если } |\{i \in \Omega : \rho(x_i, y_i) \le \varepsilon\}| \ge q_0; \\ 0, & \text{в противном случае,} \end{cases}$$

где ε и q_0 — положительные числовые параметры; $q_0 > \frac{1}{2}|\tilde{\Omega}|$.

5. Определяются оценки за класс $K_\alpha,\ \ \alpha \in \{0, 1\}$:

$$\Gamma_\alpha(\tilde{y}) = \frac{1}{k_{\ \alpha}} \sum_{\Omega \in \tilde{\Omega}} \sum_{\tilde{x} \in T_\alpha} \omega(\tilde{x}) \cdot W(\Omega) \cdot B_\Omega(\tilde{x}, \tilde{y}).$$

6. Решающее правило, полученное в результате построения методом ABO по выборке X_l и заданных параметрах $\tilde{\Theta}$ алгоритма $A(X_l, \tilde{\Theta})$ классификации или распознавания номера класса α произвольной точки $\tilde{y}$ (не обязательно содержащейся в таблице обучения), состоит

в следующем.

$$A(X_l, \tilde{\Theta})(\tilde{y}) = \alpha, \text{ если } \Gamma_\alpha(\tilde{y}) > \Gamma_{1-\alpha}(\tilde{y}) + \beta;$$

$$\textit{иначе значение } A(X_l; \tilde{\Theta})(\tilde{y}) \textit{ не определено.}$$

Здесь $\tilde{\Theta}$ обозначает всю совокупность параметров, входящих в модель *ABO*: $\tilde{\Theta} = (\tilde{\omega}, \tilde{W}, \tilde{\Omega}, \varepsilon, q_0, \tilde{k_\alpha}, \beta)$.

Подчеркнём: A обозначает метод построения алгоритма классификации (в нашем случае — *ABO*, а $A(X_l; \tilde{\Theta})$ — полученный по методу *ABO* по заданной выборке и набору параметров $\tilde{\Theta}$ конкретный алгоритм.

В тех случаях, когда для упрощения записи можно опустить использованные параметры, будем обозначать алгоритм, построенный согласно описанной модели *ABO* как $A(X_l)$ или даже $h = A(X_l) = A(X_l, \tilde{\Theta})$.

7. Задаётся *функция потерь*. В рамках данной статьи — как

$$\lambda(h, \tilde{y}) = \begin{cases} 0, & \text{если } h(\tilde{y}) = \alpha(\tilde{y}); \\ 1, & \text{если } h(\tilde{y}) \neq \alpha(\tilde{y}) \ \text{ или } h(\tilde{y}) \text{ не определено,} \end{cases} \tag{6}$$

где $h(\tilde{y})$ — номер класса, определённый построенным алгоритмом, а $\alpha(\tilde{y})$ — истинный номер класса, которому принадлежит точка $\tilde{y}$. Заданная таким образом функция потерь λ является характеристической функцией ошибки.

8. Эмпирический функционал качества

$$\frac{1}{l} \sum_{\tilde{y} \in X_l} \lambda(h, \tilde{y})$$

определяет частоту ошибок при распознавании объектов произвольной обучающей выборки X_l длины l.

9. Обучение в рамках модели *ABO* реализуется последовательной адаптивной процедурой нахождения таких значений параметров, при которых достигается как можно меньшее значение функционала качества на заданной для обучения выборки.

В классической модели *ABO* используются две выборки, которые называют обучающей и контрольной (что не всегда совпадает с соответствующими названиями выборок в других моделях). "Обучающая" выборка в модели *ABO* служит для задания эталонов, а контрольная — для обучения модели путём варьирования её параметров. При таком разбиении для обучения используется вторая, контрольная выборка, а "обучающая"(эталонная) выборка служит для фиксации конечно параметрической модели, имеющей VC размерность, которая была оценена в работах [15, 17] А. Г. Ицковым — сверху как mn, а В. В. Мамаевым — снизу как mn и сверху как $4mn$, где n — число переменных–признаков, а m — число примеров в эталонной выборке. Полученные ими результаты совместны, если только ёмкость эталонной части модели *ABO* равна в точности mn.

Но даже при использовании только эталонной выборки алгоритмы распознавания, основанные на вычислении оценок, и прежде всего — тестовые, показали практическую полезность и высокую точность, что с позиций современного представления о машинном обучении объясняется построением в этих алгоритмах кратчайших элементарных классификаторов-отделителей (осуществляется минимизация их сложности в колмогоровском смысле [7]) и бустинга [28] над ними.

Использованию в моделях *АВО* обучающей выборки "напрямую" по назначению препятствует следующий факт.

Теорема 4. *Пусть модель АВО обучается по параметрам только по одной заданной обучающей выборке длины l. Тогда для любого из 2^l возможных вариантов принадлежности обучающих примеров представленной выборки одному из двух классов в рамках этой модели может быть построен алгоритм, правильно классифицирующий все объекты обучающей выборки.*

Доказательство. Функции близости B_Ω определяют для любой точки $\tilde{y}$ некоторую специфическую окрестность в виде объединения цилиндрических множеств

$$\mathscr{O}(\tilde{y}) = \bigcup_{\Omega \in \tilde{\Omega}} \{\tilde{x} \in X_l : |\{i \in \Omega : \rho(x_i, y_i) \leq \varepsilon\}| > q_0\}.$$

Построим выборку X_l из l объектов так, что $\rho(x_i^j, x_i^q) >> \varepsilon, \;\; i \in \overline{1,n},$ для любой пары точек $\tilde{x}_j \neq \tilde{x}_q$, которые будут в эту выборку входить. Очевидно, это можно сделать, взяв в качестве точек выборки центры достаточно удалённых друг от друга шаров в $\mathbb{R}^n$.

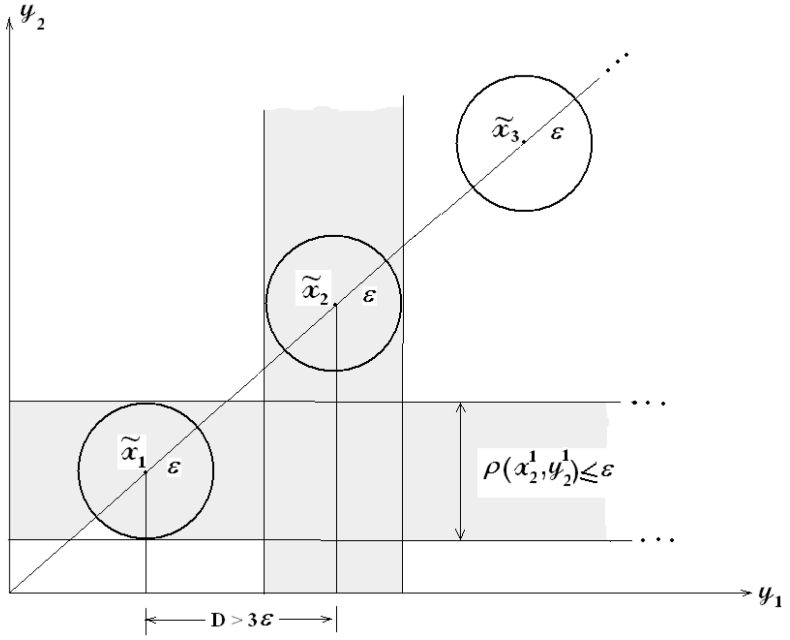

Рис. 1. Расположение шаров радиуса ε с центрами, соответствующими точкам $\tilde{x}_1, \tilde{x}_2, \ldots$ обучающей выборки.

Получить такие центры можно, например, последовательно увеличивая все координаты предыдущего центра на одну и ту же величину D, превышающую, например, $3 \cdot \varepsilon$ (рис. 1). Тогда любой точке $\tilde{x}_j$, $j = 1, \ldots, l$, обучающей выборки X_l построенный алгоритм присвоит тот же самый номер класса, который она имеет в выборке:

$$\Gamma_\alpha(\tilde{x}) = \frac{1}{k_\alpha} \sum_{\Omega \in \tilde{\Omega}} \sum_{\tilde{x} \in T_\alpha} \omega(\tilde{x}) \cdot W(\Omega) \cdot B_\Omega(\tilde{x}, \tilde{y}) = \frac{\omega(\tilde{x} \in T_\alpha)}{k_\alpha} \sum_{\Omega \in \tilde{\Omega}} W(\Omega) > \beta,$$

и в то же время $\Gamma_{1-\alpha}(\tilde{x}) = 0$. Следовательно, для любого из 2^l вариантов пометок точек выборки бинарными номерами классов можно построить алгоритм, способный настроиться на этот вариант разбиения. $\square$

Замечание 1. При настройке алгоритма (*мы принципиально различаем настройку и обучение!*) с целью получения любой заданной классификации достаточно варьировать лишь один параметр ε; остальные *разумно заданные параметры* модели не повлияют на результат такой настройки.

Напомним [24, 27], что PAC-обучаемость произвольной модели $\mathfrak{M}$ имеет место тогда и только тогда $VCD(\mathfrak{M}) < \infty$.

Следствие 2. *Модель ABO, используемая в случае одной обучающей выборки, имеет неограниченную VC-размерность: $VCD(ABO) = \infty$.*

Следствие 3. *В случае использования одной обучающей выборки, модель ABO не удовлетворяет достаточному условию обучаемости статистической теории Вапника–Червоненкиса и не является PAC обучаемой.*

PAC-обучаемость в классической модели ABO обеспечивается за счёт того, что к "обучающей" выборке добавляется так называемая "контрольная", и эта контрольная выборка используется для адаптации параметров и минимизации на ней эмпирического риска. Полагается, что m примеров "обучающей" выборки обеспечивают получение фиксированного числа m эталонов. Построенная модель с зафиксированными эталонами обладает конечной вапниковской энтропией, определяемой только пространством настраиваемых параметров. Только затем осуществляется обучение по "контрольной" выборке.

Модифицированная модель ABO^* имеет следующие отличия от описанной выше модели ABO.

1^o. Изменена область суммирования для внутренней суммы в формуле вычисления оценок:

$$\Gamma^*_\alpha(\tilde{y}) = \frac{1}{|T_\alpha| - 1} \sum_{\Omega \in \tilde{\Omega}} \sum_{\{\tilde{x} \in T_\alpha :\ \tilde{x} \neq \tilde{y}\}} \omega(\tilde{x}) \cdot W(\Omega) \cdot B_\Omega(\tilde{x}, \tilde{y}).$$

Область суммирования $\{\tilde{x} \in T_\alpha :\ \tilde{x} \neq \tilde{y}\}$ исключает вклад в оценку самой оцениваемой точки $\tilde{y}$, т.е. точка сама за себя не "голосует".

Нормирующий коэффициент $1/(|T_\alpha|-1)$ содержит в знаменателе число, в точности равное количеству точек класса T_α, по которым происходит суммирование.

Естественно полагать, что если l — число всех примеров обучающей выборки, то $|T_\alpha|$ представляет её некоторую часть, например, близкую к половине, $k_\alpha = |T_\alpha| \approx l/2$.

2^o. *Перед началом обучения все точки выборки сортируются в лексикографическом порядке.*

Эта сортировка считается внутренним *начальным фрагментом алгоритма обучения*, который сначала накапливает объекты в некоторый буфер в порядке их поступления и затем, до начала основной своей части, сортирует их. Благодаря этому любое изменение порядка примеров в изначально заданной обучающей выборке приводит к одной и той же последовательности примеров в отсортированной выборке, которая используется для обучения.

Очевидно, что *любой алгоритм обучения модели ABO^* с учётом предварительной сортировки* (2^o) *будет симметричным.*

3^o. Решающее правило классификации или распознавания номера класса α произвольной точки $\tilde{y}$ (не обязательно содержащейся в таблице обучения) состоит в следующем.

$$A(X_l,\tilde{\Theta})(\tilde{y}) = \alpha, \text{ если } \Gamma_\alpha(\tilde{y}) > \Gamma_{1-\alpha}(\tilde{y}) + \beta \ \Leftrightarrow\ \Gamma_\alpha(\tilde{y}) - \Gamma_{1-\alpha}(\tilde{y}) > \beta;$$

$$A(X_l,\tilde{\Theta})(\tilde{y}) \ \text{ не определено, если} |\Gamma_\alpha(\tilde{y}) - \Gamma_{1-\alpha}(\tilde{y})| < \beta,$$

где β — пороговое значение, заданный параметр алгоритма; $\tilde{\Theta}$ обозначает всю совокупность параметров, входящих в модель ABO^*: $\tilde{\Theta} = (\tilde{\omega}, \tilde{W}, \tilde{\Omega}, \varepsilon, q_0, \tilde{k_\alpha}, \beta)$.

4^o. Функция потерь определяется соотношением (6).

С целью получения доказательства LOO устойчивости алгоритмов, построенных в рамках модели ABO^*, *будем проводить обучение, варьируя только параметры* $\omega(\tilde{x})$, $\tilde{x} \in \{T_0 \cup T_1\}$, чего достаточно для получения требуемого теоретического результата.

Если переписать формулу вычисления оценок в виде

$$\Gamma^*_\alpha(\tilde{y}) = \sum_{\{\tilde{x}\in T_\alpha:\ \tilde{x}\neq\tilde{y}\}} \omega(\tilde{x}) \left(\frac{1}{k_\alpha - 1} \sum_{\Omega\in\tilde{\Omega}} W(\Omega)\cdot B_\Omega(\tilde{x},\tilde{y}) \right), \tag{7}$$

то можно интерпретировать выражения

$$\mathscr{K}(\tilde{x},\tilde{y}) = \frac{1}{k_\alpha - 1} \sum_{\Omega\in\tilde{\Omega}} W(\Omega)\cdot B_\Omega(\tilde{x},\tilde{y})$$

как потенциальные функции [1]; тогда

$$\Gamma^*_\alpha(\tilde{y}) = \sum_{\{\tilde{x}\in T_\alpha:\ \tilde{x}\neq\tilde{y}\}} \omega(\tilde{x})\cdot \mathscr{K}(\tilde{x},\tilde{y}),$$

и параметры параметры $\omega(\tilde{x}),\ \tilde{x} \in \{T_0 \cup T_1\}$ можно интерпретировать как коэффициенты в разложении оценки по потенциальным функциям. Заметим, что указанная интерпретация служит лишь некоторым обоснованием целесообразности выбора именно и только этих параметров для описываемой ниже процедуры обучения.

5. Алгоритм $\mathscr{A}$ обучения модифицированной модели ABO^* и его свойства

Опишем алгоритм обучения модели ABO^*, который будем далее обозначать A^*.

Чтобы упростить формулы, будем полагать, что зафиксированы веса опорных множеств: $W(\Omega) = 1$ для всех $\Omega \in \tilde{\Omega}$.

Пусть задана константа $\beta > 0$ и пусть $\omega_0 > 0$ — начальное значение, одинаковое для всех параметров $\omega(\tilde{x})$ — весов точек $\tilde{x} \in \{T_0 \cup T_1\}$, определяющее начальные значения оценок $\Gamma_\alpha^{*(0)}(\tilde{y}),\ \alpha \in \{0, 1\}$.

При предъявлении на очередном шаге обучения $t \in \{1, 2, \dots\}$ очередной точки $\tilde{y}$ обучающей выборки будем полагать, что имеет место ошибка в двух случаях.

В первом случае — если одновременно

$$\begin{cases} \Gamma_\alpha^{*(t)}(\tilde{y}) - \Gamma_{1-\alpha}^{*(t)}(\tilde{y}) > \beta, \\ \text{и точка } \tilde{y} \text{ принадлежит классу с номером } 1-\alpha\ . \end{cases} \tag{8}$$

Заметим, что при выполнении условия (8) разность $\Gamma_\alpha^{*(t)}(\tilde{y}) - \Gamma_{1-\alpha}^{*(t)}(\tilde{y})$ положительна. Для корректировки ошибки необходимо уменьшить оценку $\Gamma_\alpha^{*(t)}(\tilde{y})$ и(или) увеличить оценку $\Gamma_{1-\alpha}^{*(t)}(\tilde{y})$.

Во втором случае будем считать, что ошибка имеет место, если

$$|\Gamma_\alpha^{*(t)}(\tilde{y}) - \Gamma_{1-\alpha}^{*(t)}(\tilde{y})| \le \beta,$$

и тогда для её корректировки нужно увеличивать оценку $\Gamma_a^{*(t)}(\tilde{y})$ по тому классу $K_a,\ a \in \{0; 1\}$, которому принадлежит очередная предъявляемая точка $\tilde{y}$ обучающей выборки и/или уменьшать оценку по классу K_{1-a}.

В случае, противном указанным двум, ошибки при предъявлении очередной точки $\tilde{y}$ нет, и веса $\omega_t(\tilde{x})$ не изменяются.

Будем полагать заданными положительные константы Δ_1 и Δ_2. Обозначим $\Delta_0 = \Delta_1 + \Delta_2$. Пусть, далее, $\gamma_t = \dfrac{1}{2^t}$ есть монотонно убывающая последовательность положительных чисел, $t = 0, 1, 2, \dots$.

В случае, если при предъявлении очередного примера $\tilde{y}$, принадлежащего классу $K_{1-\alpha}$, имеет место ошибка, выполняется процедура *коррекции весов ровно* μ *точек* $\tilde{x} \in \{T_0 \cup T_1\}$, ближайших к точке $\tilde{y} = 1$ по евклидовой метрике

$$\rho(\tilde{x}, \tilde{y}) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

по формулам

$$\begin{cases} \omega_t(\tilde{x}) = \max\{\omega_{t-1}(\tilde{x}) - \gamma_t\Delta_1;\ 0\}, & \text{если } \tilde{x} \text{ принадлежит классу } \alpha, \\ \omega_t(\tilde{x}) = \omega_{t-1}(\tilde{x}) + \gamma_t\Delta_2, & \text{если } \tilde{x} \text{ принадлежит классу } 1-\alpha. \end{cases} \tag{9}$$

Если же очередной пример $\tilde{y}$ классифицируется построенным к моменту его предъявления алгоритмом правильно, то изменение веса $\omega_{t-1}(\tilde{x})$ не происходит ни для одной точки $\tilde{x}$ обучающей выборки. *Константа μ является параметром алгоритма обучения.*

Очевидно, вес $\omega_t(\tilde{x})$ каждого примера $\tilde{x}$ на каждом шаге t будет неотрицательным. Под шагом t понимается очередной номер предъявления примера обучающей выборки.

При необходимости, если номер шага t превышает заданную длину обучающей выборки l, полагается, что выборка циклически повторно просматривается сначала до конца.

Согласно процедуре (9), в результате коррекции на шаге t будем иметь

$$\gamma_t \min\{\Delta_1, \Delta_2\} \le |\omega_t(\tilde{x}) - \omega_{t-1}(\tilde{x})| \le \gamma_t \max\{\Delta_1, \Delta_2\} < \gamma_t\Delta_0.$$

Величина суммарного изменения оценки $\Gamma_\alpha^{*(t)}(\tilde{y})$ $(\alpha \in \{0;1\})$ в результате коррекции на шаге t будет заведомо ограничена величиной

$$\frac{\gamma_t \cdot \mu \cdot \Delta_0 \cdot |\tilde{\Omega}|}{k_\alpha - 1},$$

поскольку вклад в суммирование, как оговорено выше, дают только μ точек.

6. CV_{Loo} устойчивость алгоритма обучения $\mathscr{A}$

Теорема 5. *Алгоритм обучения $\mathscr{A}$ является CV_{Loo} устойчивым независимо от распределений вероятностей классифицируемых объектов и обучающих выборок.*

Доказательство. Обозначим $A(X_l)$ алгоритм классификации, полученный описанным выше модифицированным методом ABO^* по заданной выборке X_l длины l, а $A(X_l^j)$ — по той же выборке X_l, из которой удалён ровно один пример с номером $j \in \overline{1,l}$. Пусть $\alpha(\tilde{x}_j)$ — номер класса этого удаляемого объекта в таблице обучения. Функция потерь

$$\lambda\Big(A(X_l), \tilde{x}_j\Big) = \begin{cases} 0, & \text{если } A(X_l)(\tilde{x}_j) = \alpha(\tilde{x}_j); \\ 1, & \text{если } A(X_l)(\tilde{x}_j) \ne \alpha(\tilde{x}_j) \ \text{ или } A(X_l)(\tilde{x}_j) \ \text{ не определено.} \end{cases}$$

Заданная таким образом функция потерь равна нулю, если построенный алгоритм правильно классифицирует точку $\tilde{x}_j$, и единице — если алгоритм даёт ошибку классификации или отказывается от решения. Можно дать следующую интерпретацию заданной функции потерь: любой вариант ответа алгоритма, отличный от правильного, является ошибкой.

Классификация точки $\tilde{x}_j$ обученным алгоритмом определяется сравнением оценок $\Gamma_\alpha^*(\tilde{x}_j)$ и $\Gamma_{1-\alpha}^*(\tilde{x}_j)$.

После обучения по выборке X_l оценки имеют вид

$$\Gamma^*_\alpha(\tilde{x}_j) = \frac{1}{|T_\alpha| - 1} \sum_{\{\tilde{x} \in T_\alpha :\ \tilde{x} \neq \tilde{x}_j\}} \omega(\tilde{x}) \sum_{\Omega \in \tilde{\Omega}} W(\Omega) \cdot B_\Omega(\tilde{x_j}, \tilde{x}); \tag{10}$$

$$\Gamma^*_{1-\alpha}(\tilde{x}_j) = \frac{1}{|T_{1-\alpha}|} \sum_{\{\tilde{x} \in T_{1-\alpha}\}} \omega(\tilde{x}) \sum_{\Omega \in \tilde{\Omega}} W(\Omega) \cdot B_\Omega(\tilde{x_j}, \tilde{x}). \tag{11}$$

Назовём величину

$$\mathscr{W}(\tilde{y}, \tilde{x}) = \omega(\tilde{x}) \sum_{\Omega \in \tilde{\Omega}} W(\Omega) \cdot B_\Omega(\tilde{y}, \tilde{x})$$

обобщённым расстоянием от точки $\tilde{y}$ до точки $\tilde{x}$. Заметим, что по окончании обучения вес любой точки $\omega(\tilde{x})$

$$0 < \omega(\tilde{x}) \leq \omega_0 + \triangle_0 \sum_{t=1}^{\infty} \frac{1}{2^t} = \omega_0 + \triangle_0 < \infty$$

будет ограниченной величиной. Обобщённое расстояние также является ограниченным:

$$0 < \mathscr{W}(\tilde{y}, \tilde{x}) < (\omega_0 + \triangle_0)|\tilde{\Omega}|.$$

Точки обучающей выборки случайно и независимо извлекаются из генеральной совокупности $X^n \subset \mathbb{R}^n$ в соответствии с существующими условными распределениями вероятностей $P_\alpha,\ \alpha \in \{0;1\}$, точек классов K_α. Интеграл

$$M_\alpha(\tilde{x}_j) = \int\limits_{X^n} \mathscr{W}(\tilde{x}_j, \tilde{x}) dP_\alpha(\tilde{x}) < \infty$$

является математическим ожиданием обобщённого расстояния от точки $\tilde{x}_j$ по классу K_α, а величина

$$\Gamma^*_\alpha(\tilde{x}_j) = \frac{1}{|T_\alpha| - 1} \sum_{\{\tilde{x} \in T_\alpha :\ \tilde{x} \neq \tilde{x}_j\}} \mathscr{W}(\tilde{x}_j, \tilde{x}) \tag{12}$$

является выборочной средней этого же обобщённого расстояния (заметим, что в сумме в правой части равенства (12) ровно $|T_\alpha| - 1$ членов).

Дисперсия

$$D_\alpha(\tilde{x}_j) = \int\limits_{X^n} (\mathscr{W}(\tilde{x}_j, \tilde{x}) - M_\alpha(\tilde{x}_j))^2 dP_\alpha(\tilde{x}) < \big((\omega_0 + \triangle_0)|\tilde{\Omega}|\big)^2$$

также является ограниченной. Тогда по теореме Чебышева при $l \to \infty$ и $|T_\alpha| \to \infty$

$$\Gamma^*_\alpha(\tilde{x}_j) \to M_\alpha(\tilde{x}_j).$$

Напомним: мы полагаем, что при длине выборки $l \to \infty$ число примеров класса K_α в этой выборке тоже стремится к бесконечности: $|T_\alpha| \to \infty,\ \alpha \in \{0;1\}$.

Возможны только следующие два случая.

1. При сколь угодно большой длине выборки l $(l \to \infty)$ обученный алгоритм принимает решение о принадлежности точки $\tilde{x}_j$ классу K_α на основе имеющего место неравенства

$$\Gamma^*_\alpha(\tilde{x}_j) - \Gamma^*_{1-\alpha}(\tilde{x}_j) > \beta,$$

где β — константа, параметр алгоритма. При $l \to \infty$ это неравенство принимает вид

$$M_\alpha(\tilde{x}_j) - M_{1-\alpha}(\tilde{x}_j) > \beta,$$

и тогда

$$M_\alpha(\tilde{x}_j) - M_{1-\alpha}(\tilde{x}_j) = \beta + \xi,$$

где ξ — некоторая положительная константа, поскольку числа M_α, $M_{1-\alpha}$ и θ являются константами (не зависящими от l) в рамках рассматриваемой вероятностной модели.

Пусть точка $\tilde{x}_j$ обучающей выборки действительно принадлежит классу K_α (алгоритм принимает правильное решение).

После обучения алгоритмом по выборке X_l^j (по той же самой выборке X_l из которой *удалён один пример* $(\tilde{x}_j, \alpha)$ *из части выборки* T_α, *состоящей из примеров класса* K_α) оценки имеют вид

$$\Gamma^{*j}_\alpha(\tilde{x}_j) = \frac{1}{|T_\alpha| - 1} \sum_{\Omega \in \tilde{\Omega}} \sum_{\{\tilde{x} \in T_\alpha \setminus \{\tilde{x}_j\}\}} \hat{\omega}(\tilde{x}) \cdot W(\Omega) \cdot B_\Omega(\tilde{x}, \tilde{y}); \tag{13}$$

$$\Gamma^{*j}_{1-\alpha}(\tilde{x}_j) = \frac{1}{|T_{1-\alpha}|} \sum_{\Omega \in \tilde{\Omega}} \sum_{\{\tilde{x} \in T_{1-\alpha}\}} \hat{\omega}(\tilde{x}) \cdot W(\Omega) \cdot B_\Omega(\tilde{x}, \tilde{y}), \tag{14}$$

Сравним оценки (10) и (13)при $l \to \infty$. Они, вообще говоря, могут отличаться некоторыми весами ($\omega(\tilde{x})$ и $\hat{\omega}(\tilde{x})$), поскольку при обучении по выборке X_l пример $\tilde{x}_j$ участвовал в обучении и поэтому мог повлиять на процесс коррекции весов. Но при $l \to \infty$ он мог инициировать коррекцию только один раз, поэтому пример $\tilde{x}_j$ мог повлиять на веса лишь μ ближайших к нему по евклидовой метрике точек из сколь угодно большого числа l предъявленных.

Выборка X_l^j состоит точно тех же примеров, что и выборка X_l, за исключением примера $\tilde{x}_j$. При обучении по выборке X_l^j пример $\tilde{x}_j$ не участвует в коррекции, и указанные выше μ точек по этому примеру не корректируются. Поэтому суммы (10) и (13) отличаются не более чем на положительную величину $\sigma_1(\tilde{x}_j)$, которую можно оценить сверху

$$\sigma_1(\tilde{x}_j) < \frac{\mu \cdot (\omega_0 + \Delta_0) \cdot |\tilde{\Omega}|}{|T_\alpha| - 1} \to 0 \text{ при } l \to \infty,\ |T_\alpha| \to \infty.$$

Аналогично, оценки (11) и (14) не могут отличаться более чем на положительную величину $\sigma_2(\tilde{x}_j)$, которую можно оценить сверху точно также

$$\sigma_2(\tilde{x}_j) < \frac{\mu \cdot (\omega_0 + \Delta_0) \cdot |\tilde{\Omega}|}{|T_{1-\alpha}|} \to 0 \text{ при } l \to \infty,\ |T_{1-\alpha}| \to \infty;$$

$$\sigma_0(\tilde{x}_j) = \sigma_1(\tilde{x}_j) + \sigma_2(\tilde{x}_j) \to 0$$

как сумма двух бесконечно малых.

При $l \to \infty$

$$\Gamma^*_\alpha(\tilde{x}_j) - \Gamma^*_{1-\alpha}(\tilde{x}_j) > \beta \;\Leftrightarrow\; \Gamma^*_\alpha(\tilde{x}_j) - \Gamma^*_{1-\alpha}(\tilde{x}_j) = \varepsilon + \beta,$$

где $\varepsilon > 0$ — некоторая константа, $\varepsilon > \sigma_0(\tilde{x}_j)$, и

$$\Gamma^{*j}_\alpha(\tilde{x}_j) - \Gamma^{*j}_{1-\alpha}(\tilde{x}_j) = \beta + \varepsilon - \sigma_0(\tilde{x}_j) > \beta,$$

и в пределе при $l \to \infty$ алгоритмы $A(X_l)$ и $A(X_l^j)$, полученные в результате обучения по выборкам X_l и X_l^j будет давать одинаковый (в этом случае — правильный) ответ при классификации точки $\tilde{x}_j$ с вероятностью единица.

Точно таким же способом показывается, что результат классификации алгоритмами $A(X_l)$ и $A(X_l^j)$ в случае, когда, как и выше, $\Gamma^*_\alpha(\tilde{x}_j) - \Gamma^*_{1-\alpha}(\tilde{x}_j) > \beta$, но точка $\tilde{x}_j$ принадлежит классу $K_{1-\alpha}$ (в этом случае — уже ошибочный) также изменяться не будет. Таким образом, для любого $\alpha \in \{0;1\}$

$$\Gamma^*_\alpha(\tilde{x}_j) - \Gamma^*_{1-\alpha}(\tilde{x}_j) > \beta \quad \Rightarrow \Gamma^{*j}_\alpha(\tilde{x}_j) - \Gamma^{*j}_{1-\alpha}(\tilde{x}_j) > \beta.$$

Аналогичным способом легко показать обратное:

$$\Gamma^{*j}_\alpha(\tilde{x}_j) - \Gamma^{*j}_{1-\alpha}(\tilde{x}_j) > \beta \quad \Rightarrow \Gamma^*_\alpha(\tilde{x}_j) - \Gamma^*_{1-\alpha}(\tilde{x}_j) > \beta.$$

2. При сколь угодно большой длине выборки l $(l \to \infty)$ обученный алгоритм принимает решение об ошибке на основе имеющего место неравенства $|\Gamma^*_\alpha(\tilde{x}_j) - \Gamma^*_{1-\alpha}(\tilde{x}_j)| < \beta$. В этом случае

$$|\Gamma^*_\alpha(\tilde{x}_j) - \Gamma^*_{1-\alpha}(\tilde{x}_j)| = \beta - \varepsilon; \quad \varepsilon > 0;$$

$$|\Gamma^{*j}_\alpha(\tilde{x}_j) - \Gamma^{*j}_{1-\alpha}(\tilde{x}_j)| < \beta - \varepsilon + \sigma_0(\tilde{x}_j),\ \varepsilon = const > \sigma_0(\tilde{x}_j) \to 0 \text{ при } l \to \infty,$$

поэтому $\Gamma^{*j}_\alpha(\tilde{x}_j) - \Gamma^{*j}_{1-\alpha}(\tilde{x}_j)| < \beta$, и в этом случае результат классификации точки $\tilde{x}_j$ алгоритмами $A(X_l)$ и $A(X_l^j)$ в пределе при $l \to \infty$ будет одинаковым.

В силу доказанного факта неизменяемости результатов классификации алгоритмами $A(X_l)$ и $A(X_l^j)$, при $l \to \infty$ получаем

$$\forall j \in \{1, \dots, l\} \;\; P^l\{|\lambda(A(X_l^j), \tilde{x}_j)) - \lambda(A(X_l), \tilde{x}_j))| = 0\} = 1,$$

откуда немедленно следует CV_{Loo} устойчивость обучающего алгоритма $\mathscr{A}$. □

7. $ELoo_{err}$ и LOO устойчивость и обучаемость модифицированной модели ABO^*

Теорема 6. *Алгоритм $\mathscr{A}$ обучения модели ABO^* с функцией потерь (6) является $ELoo_{err}$ устойчивым независимо от вида распределений вероятностей классифицируемых объектов и обучающих выборок.*

Доказательство. Рассмотрим математическое ожидание

$$\mathbf{E}(A(X_l)) = \int\limits_{X^n\times\{0;1\}} \lambda\big(A(X_l)(\tilde{x}), (\tilde{x}, \alpha)\big) d\hat{P}(\tilde{x}, \alpha), \tag{15}$$

где $\hat{P}$ — вероятностная мера на множестве $X^n \times \{0; 1\}$, которое является вероятностью события, определяемого этой моделью как *ошибка алгоритма, обученного по произвольной заданной выборке длины* l, и включающего в себя как неправильную классификацию точек, так и отказ от классификации. Иначе говоря, величина (15) есть вероятность ошибки алгоритма $A(X_l)$.

Статистика

$$\frac{1}{l}\sum_{j=1}^{l} \lambda(A(X_l^j), \tilde{x}_j))$$

является выборочной средней или частотой ошибки алгоритма $A(X_l)$, подсчитываемой методом скользящего контроля, и являющейся при $l \to \infty$ почти несмещённой оценкой [4, с. 267] величины $\mathbf{E}(A(X_l))$ в том смысле, что если обозначить

$$\mathbf{E}(A(X_l)) = \bar{\mathbf{p}}_l,$$

то математическое ожидание оценки, полученной методом скользящего контроля будет

$$\mathbf{E}\Big(\frac{1}{l}\sum_{j=1}^{l} \lambda(A(X_l^j), \tilde{x}_j))\Big) = \bar{\mathbf{p}}_{l-1}.$$

Величины $\bar{\mathbf{p}}_{l-1}$ и $\bar{\mathbf{p}}_l$ стремятся при $l \to \infty$ к одному и тому же пределу. Обозначим этот предел $\bar{\mathbf{p}}_\infty$. При $l \to \infty$ разность $\bar{\mathbf{p}}_{l-1} - \bar{\mathbf{p}}_\infty = \epsilon_1(l)$, где $\epsilon_1(l)$ — бесконечно малая, которая независимо от её знака стремится к нулю; $\epsilon_1(l) = o(l)$. Аналогично $\bar{\mathbf{p}}_l - \bar{\mathbf{p}}_\infty = \epsilon_2(l)$, $\epsilon_2(l) = o(l)$. Поэтому

$$|\bar{\mathbf{p}}_l - \bar{\mathbf{p}}_{l-1}| = |\bar{\mathbf{p}}_\infty + \epsilon_2(l) - \bar{\mathbf{p}}_\infty - \epsilon_1(l)| \to 0 \ \text{ при } l \to \infty.$$

Следовательно, при $l \to \infty$, согласно закону больших чисел,

$$\bar{\mathbf{p}}_{l-1} = \frac{1}{l}\sum_{j=1}^{l} \lambda(A(X_l^j), \tilde{x}_j)) \to \mathbf{E}(A(X_l)) = \bar{\mathbf{p}}_l.$$

Таким образом, для любой выборки X_l и для любого $j \in \{0, \dots, l\}$ имеет место неравенство

$$|\mathbf{E}(A(X_l)) - \frac{1}{l}\sum_{j=1}^{l} \lambda(A(X_l^j), \tilde{x}_j))| \to 0.$$

Последнее соотношение означает, что

$$\forall \varepsilon \quad \exists l^0 = l^0(\varepsilon): \quad \forall l > l^0 \quad |\mathbf{E}(A(X_l)) - \frac{1}{l}\sum_{j=1}^{l} \lambda(A(X_l^j), \tilde{x}_j))| < \varepsilon. \tag{16}$$

Иначе говоря, в пределе при $l \to \infty$ неравенство (16) выполняется достоверно, с вероятностью единица:

$$P^l\{|\mathbf{E}(A(X_l)) - \frac{1}{l}\sum_{j=1}^{l} \lambda(A(X_l^j), \tilde{x}_j))| \leq \epsilon(l)\} = 1$$

для любой бесконечно малой $\epsilon(l) \to 0$ при $l \to \infty$; и тем более для любой сколь угодно малой $\delta(l) \to 0$ и для любого $j \in \{0, \dots, l\}$ выполняется неравенство

$$P^l\{|\mathbf{E}(A(X_l)) - \frac{1}{l}\sum_{j=1}^{l} \lambda(A(X_l^j), \tilde{x}_j))| \leq \epsilon(l)\} \geq 1 - \delta(l),$$

означающее $ELoo_{err}$ устойчивость алгоритма $\mathscr{A}$. □

Теорема 7. *Существует LOO устойчивый алгоритм обучения модифицированной модели* ABO^* *методом минимизации эмпирического риска по одной заданной обучающей выборке, обеспечивающий обучаемость этой модели в форме универсального эмпирического обобщения.*

Доказательство. Описанный выше (параграф 5) алгоритм $\mathscr{A}$ обучения модели ABO^* является симметричным (стр. 36), CV_{Loo} устойчивым (по теореме 5) и $ELoo_{err}$ устойчивым (по теореме 6), следовательно, по теореме 3 обеспечивает универсальное эмпирическое обобщение по методу минимизации эмпирического риска по одной (без использования эталонной) обучающей выборке. □

Заключение

В статье получен следующий теоретический результат: существует устойчивый алгоритм $\mathscr{A}$ обучения модифицированной модели ABO^*, гарантирующий её обучаемость в форме универсального эмпирического обобщения непосредственно по единственной заданной выборке путём минимизации эмпирического риска. Для получения этого результата была доказана LOO устойчивость алгоритма $\mathscr{A}$. Алгоритм $\mathscr{A}$ подробно описан в данной статье и является процедурой обучения с поощрением. Он предполагает адаптацию только весов объектов обучающей выборки. Остальные параметры модели полагаются зафиксированными. Этого оказалось достаточно для достижения требуемого результата.

Предложенная модификация модели ABO является минимальной: исключается только случай, когда "точка голосует сама за себя".

Нетрудно показать, что в случае, когда в модифицированной модели ABO^* процесс обучения заключается только в выборе по обучающей выборке кратчайших элементарных логических отделителей (в частности — тупиковых тестов), универсальное эмпирическое обобщение также будет иметь место.

Список литературы

1. *Айзерман М. А., Браверман Э. М., Розоноэр Л. И.* Метод потенциальных функций в теории обучения машин / М. А. Айзерман, Э. М. Браверман, Л. И. Розоноэр. — М.: Наука, 1970. — 320 с.

   AIZERMAN, M. A., BRAVERMAN, E. M., & ROZONOER, L. I. (1970) *Method of potential functions in the theory of learning machines.* Moscow: Nauka.

2. *Ашуров А. Р., Рудаков К. В.* Алгоритмы вычисления оценок для задачи распознавания объектов с континуальной начальной информацией / А. Р. Ашуров, К. В. Рудаков // Ж. вычисл. матем. и матем. физ., 1984. — Т 24.— №12. — С. 1871–1880.

   ASHUROV, A. R., RUDAKOV, K. V. (1984) Algorithms of Estimates Calculation for the Recognition Problem whith Continued Initial Information. *J. Comput. Math.* & *Math. Phisycs*, 24 (12), p. 1871–1880.

3. *Вапник В. Н., Червоненкис А. Я.* Теория распознавания образов. М.: Наука, 1974. — 416 с.

   VAPNIK, V. N., CHERVONENKIS, A. Ya. (1974) *Pattern Recognition Theory.* Moscow: Nauka.

4. *Вапник В. Н.* Восстановление зависимостей по эмпирическим данным В. Н. Вапник. М.: Наука, 1979. — 448 с.

   VAPNIK, V. N. (1974) *Dependence Recovery by Use Empirical Data.* Moscow: Nauka.

5. *Воронцов К. В.* Обзор современных исследований по проблеме качества обучения алгоритмов // Таврический вестник информатики и математики, 2004. — № 1. — С. 5–24.

   VORONTZOV, K. V. (2004) Modern State of Art of the Learning Algorithms Quality. *Tavrida J. of Computer Sci. Theory and Math.*, 1, p. 5–24.

6. *Дмитриев А. Н., Журавлёв Ю. И., Кренделев Ф. П.* О математических принципах классификации предметов и явлений // Дискретный анализ, 1966. — Вып. 7. — С. 3–15.

   DMITRIEV, A. N., ZHURAVLEV Yu. I. & KRENDELEV F. P. (1966) On Mathematical Principles of Objects and Phenomenons Classification. *Descreet Analisys*, 7, p. 3–15.

7. *Донской В. И.* Колмогоровская сложность и ее применение в машинном обучении // Таврический вестник информатики и математики, 2012. — № 2. —– С. 4–35.

   DONSKOY, V. I. (2012) Kolmogorov Complexity and its Application in Machine Learning. *Tavrida J. of Computer Sci. Theory and Math.*, 2, p. 4–35.

8. *Дьяконов А. Г.* Анализ данных, обучение по прецедентам, логические игры, системы WEKA, RapidMiner и MatLab (Практикум на ЭВМ кафедры математических основ прогнозирования) / А. Г. Дьяконов. — М.: ВМК МГУ, 2010. — 278 с.

   DYACONOV, A. G. (2010) *Data Analisys, Learning by Precedents, Logical Games, WEKA, RapidMiner and MatLab systems (Practical Work on Computer, Chair of Mathematical Bases of Forecasting).* Moscow: MSU.

9. *Дьяконов А. Г.* О выборе системы опорных множеств для эффективной реализации алгоритмов распознавания типа вычисления оценок // Ж. вычисл. матем. и матем. физ., 2000. — Т 40.—– № 7. — С. 1104–1118.

DYACONOV, A. G. (2000) The Choice of Support Set System for Effective Realization of the Recognition Estimates Calculation Algorithms. *J. Comput. Math. & Math. Phisycs*, 40 (7), p. 1104–1118.

10. *Дьяконов А. Г.* Теория систем эквивалентностей для описания алгебраических замыканий обобщенной модели вычисления оценок // Ж. вычисл. матем. и матем. физ., 2011. — Т 51. — № 3. — С. 529–544.

    DYACONOV, A. G. (2011) The Equivalence Systems Theory for the describing of algebraic Closings of Generalized Estimates Calculation Model. *J. Comput. Math. & Math. Phisycs*, 51 (3), p. 529–544.

11. *Журавлев Ю. И.* Алгоритмы распознавания, основанные на вычислении оценок / Ю. И. Журавлев, В. В. Никифоров // Кибернетика, 1971. — № 3. — С. 1–11.

    ZHURAVLEV, Yu. I. (1971) Recognition Algoritms Based on Estimates Calculation. *Cibernetics*, 3, p. 1–11.

12. *Журавлев Ю. И.* Избранные научные труды / Ю. И. Журавлёв. — М.: Магистр, 1998. — 420 с.

    ZHURAVLEV, Yu. I. (1998) *Selected Scientific Works.* Moskow: Magister.

13. *Журавлев Ю. И.* Об алгебраическом подходе к решению задач распознавания или классификации / Ю. И. Журавлев // Проблемы кибернетики, 1978. — Вып. 7. — С. 5–68.

    ZHURAVLEV, Yu. I. (1978) The Algebraic Approach to Recognition or Classification Problems. *Problems of Cybernetics*, 7, p. 5–68.

14. *Заде Л. А.* Понятие лингвистической переменной и его применение к принятию приближенных решений / Л. А. Заде. — М.: Мир, 1976. — 165 с.

    ZADE, L. (1976) *The concept of a linguistic variable and its application to the adoption of approximate solutions.* Moscow: Mir.

15. *Ицков А. Г.* О емкости модели распознающих алгоритмов вычисления оценок / А. Г. Ицков // Журнал вычислительной математики и математической физики, 1982 — Т. 22.— № 4. — с. 975–981.

    IZKOV, A. G. (1982) On Capacity of the Model of Recognition Algorithms Based on Estimates Calculation. *J. Comput. Math. & Math. Phisycs*, 22 (4), p. 975–981.

16. *Камилов М. М., Мирзаев Н. М., Раджабов С. С.* Об одной модификации модели алгоритмов распознавания, основанных на вычислении оценок / М. М. Камилов, Н. М. Мирзаев, С. С. Раджабов // Доклады академии наук Республики Узбекистан, Ташкент: 2009. — Т. 258. — № 2. — с. 18–20.

    KAMILOV, M. M., Mirzaev N. M., & Radgabov S. S. (2009) On the Modification of Recognition Algorithms Based on Estimates Calculation. *Repotrs of Uzbekistan Academy of Science*, 258 (2), p. 18–20.

17. *Мамаев В. В.* Верхняя и нижняя границы емкости модели алгоритмов вычисления оценок / В. В. Мамаев // Депонир. в ВИНИТИ, 1997. — № 216-В97, 27.01.1997 г. — 5 с.

MAMAEV, V. V. (1997) Upper and Lower Bounds of the Capacity of the Model of Recognition Algorithms Based on Estimates Calculation. *Deposited whith All-Russian Institute of Scientific and Technical Information*, No. 216-B97.

18. *Матросов В. Л.* О критериях полноты модели алгоритмов вычисления оценок и её алгебраических замыканий / В. Л. Маторосов // Доклады академии наук СССР, 1981 — Т. 258. — № 4. — с. 791–796.

    MATROSOV, V. L. (1981) On the criteria of completeness of the Model of Algorithms Based on Estimates Calculation and its Algebraic Closures. *Reports of Russian Academy of Sciences*, 258 (4), p. 791–796.

19. *Матросов В. Л.* Ёмкость алгебраических расширений модели алгоритмов вычисления оценок / В. Л. Матросов // Журнал вычислительной математики и математической физики, 1984 — Т. 11. — № 5. — с. 1719–1730.

    MATROSOV, V. L. (1984) The Capacity of the Algebraic Extensions of Algorithms of Calculating Estimates. *J. Comput. Math.* & *Math. Phisycs*, 11 (5), p. 1719–1730.

20. *Плохонина Т. В.* О некорректности алгебраического замыкания второй степени семейства алгоритмов вычисления оценок / Т. В. Плохонина // Журнал вычислительной математики и математической физики, 1985 — Т. 25.— № 7. — с. 1073–1086.

    PLOKHONINA, T. V. (1985) On the Impropriety of Algebraic Closures of Second-Degree Family of Algorithms of Calculating Estimates. *J. Comput. Math.* & *Math. Phisycs*, 25 (7), p. 1073–1086.

21. *Рудаков К. В.* Полнота и универсальные ограничения в проблеме коррекции эвристических алгоритмов классификации / К. В. Рудаков // Кибернетика, 1987. — № 3. — с. 106–109.

    RUDAKOV, K. V. (1987) Completeness and Universal Constraints in the Correction Problem of Heuristic Classification Algorithms. *Cybernetics*, 3, p. 106–109.

22. *Рязанов В. В.* Оптимизация алгоритмов вычисления оценок по параметрам, характеризующим представительность эталонных строк / В. В. Рязанов // Журнал вычислительной математики и математической физики, 1976 — Т. 16. — № 6. — с. 1559–1570.

    RYAZANOV, V. V. (1976) Optimization of Algorithms of Calculating Estimates by the Parameters Characterizing the Representativity of the Reference Lines. *J. Comput. Math.* & *Math. Phisycs*, 16 (6), p. 1559–1570.

23. *Хилков А. В.* Формулы вычисления оценок для алгоритмов распознавания с опорными множествами / А. В. Хилков // Журнал вычислительной математики и математической физики, 1989 — Т. 29. — № 10. — с. 1565–1571.

    KHILKOV, A. V. The Formulas for Calculating Estimates for Recognition Algorithms with Support Sets. *J. Comput. Math.* & *Math. Phisycs*, 29 (10), p. 1565–1571.

24. BLUMER A., EHRENFEUCHT A., HAUSSLER D., and WARMUTH M. K. Learnability and the Vapnik-Chervonenkis dimension / A. Blumer, A. Ehrenfeucht, D. Haussler and M. K. Warmuth // Journal of the ACM, 1989. — 36. — № 4.— p. 929–865.

25. BOUSQUET O., ELISSEEFF A. Algorithmic Stability and Generalization Performance / Olivier Bousquet , Andre Elisseeff // Advances in Neural Information Processing Systems, 2001. — № 13. — P. 196–202.

26. MUKHERJEE S. Learning theory: stability is sufficient for generalization and necessary and sufficient for consistency of empirical risk minimization / Sayan Mukherjee, Partha Niyogi, Tomaso Poggio and Ryan Rifkin // Advances in Computational Mathematics, 2006. — No 25. — p. 161–193.

27. PESTOV V. PAC learnability versus VC dimension: a footnote to a basic result of statistical learning / Vladimir Pestov // Proc. 2011 International Joint Confernce on Neural Networks (IJCNN'2011), San Jose, CA (July 30 – Aug. 5, 2011), p. 1141–1145.

28. SCHAPIRE R. E.The Strength of Weak Learnability / Robert E. Schapire // Machine Learning, 1990. — № 5. — p. 197–227.

УДК: 519.87

MSC2010: 90C47

ЛИНЕЙНАЯ ЗАДАЧА ПРИНЯТИЯ РЕШЕНИЙ С УЧЕТОМ РИСКОВ И СОЖАЛЕНИЙ

© А. Е. Бардин, Ю. Н. Житенева

Государственный гуманитарно-технологический университет

факультет информатики

ул. Зеленая, 22, г. Орехово-Зуево, Московская обл., 142611, Российская Федерация

e-mail: *inth2006@rambler.ru, ulya_zhiteneva@mail.ru*

The linear problem of making solution with consider the risks and regrets.

Bardin A. E., Zhiteneva J. N.

Abstract. This paper deals with a linear programming problem with interval function coefficients. A new approach for a problem under uncertainty is proposed. It is based on two concepts: the guaranteed result and the minimax regret criterion. We consider the following linear programming problem under uncertainty

$$\max_{x\in X} f(x,y),$$

where $f(x,y)=\sum\limits_{i=1}^{n} y_i x_i$ is utility function, the feasible set X is given by

$$X=\big\{x=(x_1,x_2,\ldots,x_n)\in\mathbb{R}^n|\ Ax\le b\big\}.$$

Here A is an $m\times n$ matrix, x and b are n- and m-column vectors, respectively. The set of uncertainties is defined by

$$Y=\big\{y=(y_1,y_2,\ldots,y_n)|\ a_i\le y_i\le b_i,\ i\in\{1,2,\ldots,n\}\big\}.$$

Further, we decide the standard linear programming problem

$$f(x,y)=\sum_{i=1}^{n} a_i x_i\to\max,$$
$$\begin{cases} x=(x_1,x_2,\ldots,x_n)\in X,\\ x_i\ge 0,\ i\in\{1,2,\ldots,n\},\end{cases}$$

where $X=\big\{x=(x_1,x_2,\ldots,x_n)\in\mathbb{R}^n|\ Ax\le b\big\}$ and $x_V^*\in X$ is solution.

The two-criteria problem

$$\Gamma=\langle X,\big\{f_1(x)=R_V(x),\ f_2(x)=R_S(x)\big\}\rangle,$$

is formalized. Here

$$R_V(x)=f_V[x_V^*]-\sum_{i=1}^{n} a_i x_i$$

is risk function,

$$R_S(x) = \max_{y \in Y} \phi(x, y) - \min_{x \in X} \max_{y \in Y} \phi(x, y)$$

is regret function, where $\phi(x, y) = \max_{z \in X} f(x, y) - f(x, y)$.

Then we look for vector $x_P \in X$, which minimizes the function

$$F(x) = R_V^2(x) + R_S^2(x),\ x \in X.$$

Solution $x_P \in X$ is Pareto optimal for problem Γ. In this paper we construct the algorithm finding optimal solution $x_P \in X$. In order to illustrate the proposed solution method, a numerical example is given.

Keywords: *linear programming problem, uncertainty, two-criteria problem, risk function, regret function, Pareto minimum.*

Введение

В работах [1]–[3] Жуковского В. И. и его учеников исследуются конфликтные системы при неопределенности. Там же формализованы различные понятия решения, которые базируются на принципах оптимального поведения из теории задач при неопределенности. В данной работе анализируется линейная модель принятия решений в условиях действия неконтролируемых факторов.

Для принятия решения в условиях неопределенности используется ряд критериев. Согласно максиминному критерию Вальда игра с природой рассматривается как конфликт с разумным и агрессивным противником, делающим все для того, чтобы помешать лицу, принимающему решение (ЛПР) достигнуть успеха. Таким образом, осторожная (максиминная) стратегия ЛПР является оптимальной согласно этому критерию. Этот критерий олицетворяет «позицию пессимизма». Он ориентируется на самую неблагоприятную для ЛПР реализацию неопределенности. Такой подход естественен для того, кто боится проиграть.

Критерий минимаксного сожаления Сэвиджа при выборе оптимальной стратегии ориентируется не на выигрыш, а на сожаления. В качестве оптимальной стратегии выбирается та стратегия, при которой величина сожаления в наихудших условиях минимальна. Такая стратегия поведения часто соответствует принципам азартного игрока.

Очевидно, разумный игрок должен учитывать как возникающие риски при принятии решения, так и возможность получения большего выигрыша. Возникает идея формализации нового подхода к принятию решений в игре с природой, который мог бы соединить позитивные особенности обоих принципов и ослабить их негативные свойства. В данной работе формализуется понятие U-оптимального по рискам и сожалениям решения линейной задачи и указан алгоритм построения оптимального решения.

1. Постановка задачи. Формализация оптимального решения

Рассмотрим задачу линейного программирования при неопределенности

$$f(x,y) = \sum_{i=1}^{n} y_i x_i \to \max,$$

$$\begin{cases} x = (x_1, x_2, \dots, x_n) \in X, \ X \in comp\, \mathbb{R}^n, \\ 0 < a_i \le y_i \le b_i, \\ x_i \ge 0, \ i \in \{1, 2, \dots, n\}. \end{cases} \tag{1}$$

Неотрицательный набор переменных $x = (x_1, x_2, \dots, x_n)$ можно интерпретировать как план производства n видов продукции, причем количество продукции каждого вида ограничено ресурсами данного производства. Будем предполагать, что множество X допустимых планов определено системой линейных уравнений и нестрогих линейных неравенств. Неопределенность y_i, $i \in \{1, 2, \dots, n\}$, можно рассматривать как доход от реализации единицы i-го вида продукции. При этом для указанных доходов известны лишь граничные значения, которые зависят от рыночных цен, спроса и т.д. Тогда целевая функция $f(x,y) = \sum\limits_{i=1}^{n} y_i x_i$ есть суммарный доход от реализации всей продукции. Формализуем U-оптимальное по рискам и сожалениям решение данной задачи.

Согласно подходу в работах [1]–[3] стратегический риск по Вальду для лица принимающего решение (ЛПР) положим равным

$$R_V(x) = \max_{x \in X} \min_{y \in Y} f(x,y) - \min_{y \in Y} f(x,y), \tag{2}$$

стратегическое сожаление по Сэвиджу определим как

$$R_S(x) = \max_{y \in Y} \phi(x,y) - \min_{x \in X} \max_{y \in Y} \phi(x,y), \tag{3}$$

где функция сожаления $\phi(x,y)$ вычисляется согласно равенству

$$\phi(x,y) = \max_{z \in X} f(x,y) - f(x,y).$$

Отметим, что риски и сожаления определены как стратегические, то есть являются функциями только от выбранного решения (стратегии) $x \in X$. Заметим также, что вследствие непрерывности функции $f(x,y)$ и компактности множеств решений X

и неопределенностей

$$Y = \{y = (y_1, y_2, \ldots, y_n) |\ a_i \le y_i \le b_i,\ i \in \{1, 2, \ldots, n\}\},$$

все указанные максимумы и минимумы существуют, причем функции $R_V(x)$ и $R_S(x)$ будут непрерывными.

Для исходной задачи (1) введем обозначение:

$$f_V[x] = \min_{y \in Y} f(x, y) = \sum_{i=1}^{n} a_i x_i.$$

Тогда гарантированный суммарный доход от реализации плана x_V^* равен

$$\max_{x \in X} \min_{y \in Y} f(x, y) = \max_{x \in X} \sum_{i=1}^{n} a_i x_i = f_V[x_V^*].$$

Здесь осторожный по Вальду план $x_V^* \in X$ есть решение оптимизационной задачи

$$f(x, y) = \sum_{i=1}^{n} a_i x_i \to \max,$$
$$\begin{cases} x = (x_1, x_2, \ldots, x_n) \in X, \\ x_i \ge 0,\ i \in \{1, 2, \ldots, n\}. \end{cases}$$

Получаем стратегический риск по Вальду

$$R_V(x) = f_V[x_V^*] - \sum_{i=1}^{n} a_i x_i.$$

Далее построим алгоритм нахождения функции $R_S(x)$, определенной в (3).

Введем обозначение:

$$\Phi_S[x] = \max_{y \in Y} \phi(x, y). \tag{4}$$

Пусть

$$\phi(x, y) = \max_{z \in X} f(x, y) - f(x, y) = \max_{z \in X} \sum_{i=1}^{n} y_i (z_i - x_i) = \sum_{i=1}^{n} y_i (z_i^*(y) - x_i). \tag{5}$$

Здесь вектор $z^*(y) = (z_1^*(y), z_2^*(y), \ldots, z_n^*(y))$ является решением задачи параметрического программирования

$$f(z, y) = \sum_{i=1}^{n} y_i z_i \to \max,$$
$$\begin{cases} z = (z_1, z_2, \ldots, z_n) \in X,\ X \in comp\mathbb{R}^n \\ 0 < a_i \le y_i \le b_i, \\ z_i \ge 0,\ i \in \{1, 2, \ldots, n\}. \end{cases} \tag{6}$$

Так как параметры y_i, $i \in \{1, 2, \ldots, n\}$, положительны, а также $z_i \geq 0$, $i \in \{1, 2, \ldots, n\}$, то оптимальные решения $z^*(y)$ принадлежат множеству X^P максимальных по Парето точек множества X. Именно, точка $x^P = (x_1^P, x_2^P, \ldots, x_n^P)$ будет максимальной по Парето точкой множества, если для всех $x = (x_1, x_2, \ldots, x_n)$ несовместна система неравенств

$$\begin{cases} x_1 \geq x_1^P, \\ x_2 \geq x_2^P, \\ \ldots \\ x_n \geq x_n^P, \end{cases}$$

причем хотя бы одно из неравенств является строгим.

Таким образом, с одной стороны, для каждого фиксированного набора $y = (y_1, y_2, \ldots, y_n) \in Y$ задача (6) является задачей линейного программирования. Следовательно, оптимальное решение $z^*(y)$ достигается в некоторой угловой точке множества X^P . С другой стороны, функция $\phi(x, y)$, заданная в (5), является кусочно-линейной по переменной y. Поэтому при фиксированном наборе $x = (x_1, x_2, \ldots, x_n) \in X$ наибольшее сожаление ЛПР достигается в одной из угловых точек многогранника Y, где совокупность всех угловых точек представляет собой множество

$$Y_C = \big\{c = (c_1, c_2, \ldots, c_n) |\ c_i \in \{a_i, b_i\},\ i \in \{1, 2, \ldots, n\}\big\}.$$

Исходя из вышесказанного, получаем, что для построения явного вида функции $\Phi_S[x]$, определенной в (4), можно «сузить» множество допустимых решений X до множества X_C^P, а множество неопределенностей Y до множества Y_C. Таким образом, имеем следующий алгоритм:

1 шаг. Находим совокупность X_C^P всех угловых точек множества X^P.

2 шаг. Для каждого набора $c = (c_1, c_2, \ldots, c_n) \in Y_C$ решаем оптимизационную задачу

$$\begin{gathered} f(x) = \sum_{i=1}^{n} c_i x_i \to \max, \\ \begin{cases} c = (c_1, c_2, \ldots, c_n) \in Y_C, \\ x = (x_1, x_2, \ldots, x_n) \in X_C^P. \end{cases} \end{gathered} \tag{7}$$

3 шаг. Пусть $x^*(c) = (x_1^*(c), x_2^*(c), \ldots, x_n^*(c))$ оптимальное решение задачи (7) с заданным набором $c = (c_1, c_2, \ldots, c_n) \in Y_C$, тогда составляем функцию

$$\phi(x, c) = \sum_{i=1}^{n} c_i (x_i^*(c) - x_i).$$

4 шаг. Для каждого набора $c = (c_1, c_2, \ldots, c_n) \in Y_C$ определяем множество $X(c)$, которое задается системой неравенств

$$\phi(x, c) \geq \phi(x, d), \ \forall d \in Y_C.$$

5 шаг. Получаем явный вид функции $\Phi_S[x] = \phi(x, c)$, $x \in X(c) \cap X$, $c \in Y_C$.

В результате указанной выше последовательности шагов получаем представление множества X как объединение конечного числа подмножеств, именно,

$$X = \bigcup_{c \in Y_C} \big(X(c) \cap X\big).$$

На каждом подмножестве вида $X(c) \cap X$ функция $\Phi_S[x]$, заданная равенством (4), имеет «особый» аналитический вид, то есть

$$\Phi_S[x] = \phi(x, c) = \sum_{i=1}^{n} c_i(x_i^*(c) - x_i), \ \phi(x, c) \geq \phi(x, d), \ \forall d \in Y_C.$$

Отметим, что некоторые множества вида $X(c) \cap X$, $c \in Y_C$, могут быть пустыми.

6 шаг. Далее решаем классические задачи линейного программирования, число которых не более 2^n. Именно,

$$\begin{aligned} \Phi_S[x] = \sum_{i=1}^{n} c_i(x_i^*(c) - x_i) \to \min, \\ x \in \big(X(c) \cap X\big). \end{aligned} \tag{8}$$

7 шаг. Окончательно, сравнивая решения задач вида (8), находим «глобальный» минимум функции $\Phi_S[x]$ на множестве X. Пусть план x_S^* является оптимальным решением исходной задачи (1) согласно принципу Сэвиджа, то есть

$$\Phi_S[x_S^*] = \min_{x \in X} \max_{y \in Y} \phi(x, y) = \min_{x \in X} \Phi_S[x].$$

8 шаг. Если $c \in Y_C$, то функция $R_s(x)$ на множестве $X(c) \cap X$ имеет вид:

$$R_S(x) = \Phi_S[x] - \Phi_S[x_S^*] = \sum_{i=1}^{n} c_i(x_i^*(c) - x_i) - \Phi_S[x_S^*].$$

Перейдем к формализации U-оптимального решения для задачи (1). Исходной задаче поставим в соответствие двухкритериальную задачу

$$\Gamma = \langle X, \big\{ f_1(x) = R_V(x), \ f_2(x) = R_S(x) \big\} \rangle,$$

где ЛПР стремится получить значения каждого из критериев как можно меньше. Формализуем понятие оптимального решения в исходной задаче, используя известные понятия оптимальности для многокритериальных задач.

Определение 1. Решение в задаче (1) назовем U-*оптимальным по рискам и сожалениям*, если оно является минимальным для функции

$$F(x) = R_V^2(x) + R_S^2(x),\ x \in X.$$

Теорема 1. *Если система ограничений в задаче (1) является непустым компактом пространства $\mathbb{R}^n$, то существует U-оптимальное по рискам и сожалениям решение $x_U \in X$, которое является минимальным по Парето для задачи Γ. При этом решение x_U принадлежит множеству X^P максимальных по Парето точек множества X.*

2. Задача составления оптимального плана производства

Проиллюстрируем алгоритм нахождения U-оптимального по рискам и сожалениям решения на конкретной модели составления оптимального плана производства.

Пример. Рассмотрим задачу линейного программирования при неопределенности

$$f(x,y) = y_1x_1 + y_2x_2 \to \max,$$
$$\begin{cases} x_1 + x_2 \le 10, \\ 0 \le x_1 \le 8, \\ 0 \le x_2 \le 6, \end{cases} \tag{9}$$

где значения неопределенностей удовлетворяют неравенствам

$$1 \le y_1 \le 8,$$
$$3 \le y_2 \le 6.$$

Неотрицательный набор переменных $x = (x_1, x_2)$ можно интерпретировать как план производства двух видов продукции, причем количество продукции первого вида требуется выпустить не больше 8 единиц, второго – не более 6, а общее количество выпускаемой продукции не должно превосходить 10 единиц. Неопределенности y_1, y_2 представляют собой возможные доходы от реализации единиц соответствующих видов продукции. При этом для указанных доходов известны лишь граничные значения, которые зависят от рыночных цен, спроса и т.д. Тогда целевая функция $f(x,y) = y_1x_1 + y_2x_2$ есть суммарный доход от реализации всей продукции, где $y = (y_1, y_2)$.

Найдем U-оптимальное по рискам и сожалениям решение данной задачи.

Вычислим стратегический риск по Вальду $R_V(x)$, определенный в (2). Так как переменные x_1, x_2 и неопределенности y_1, y_2 в задаче (9) неотрицательны, то минимальное значение целевой функции $f(x,y)$ достигается при наихудших для ЛПР значениях неопределенностей, то есть

$$\min_{y\in Y} f(x,y) = f(x,y^{\min}) = x_1 + 3x_2,$$

где $y^{\min} = (1,3)$.

Для нахождения $\max\limits_{x\in X}\min\limits_{y\in Y} f(x,y)$ решаем задачу линейного программирования

$$\begin{gathered} f(x) = x_1 + 3x_2 \to \max, \\ \begin{cases} x_1 + x_2 \le 10, \\ 0 \le x_1 \le 8, \\ 0 \le x_2 \le 6. \end{cases} \end{gathered} \tag{10}$$

Оптимальное решение задачи (10) равно $x^{\max} = (x_1^{\max}, x_2^{\max}) = (4,6)$. Найденное осторожное решение является максиминным для задачи (9), при этом $\max\limits_{x\in X}\min\limits_{y\in Y} f(x,y) = 22$.

Если набор $x = (x_1, x_2)$ произвольное допустимое решение задачи, то стратегический риск по Вальду для ЛПР равен

$$R_V(x) = 22 - x_1 - 3x_2. \tag{11}$$

Перейдем к нахождению стратегического сожаления по Сэвиджу $R_S(x)$, определенного в (3). Используем приведенный в работе алгоритм.

1 шаг. Совокупность всех угловых точек множества X^P есть множество

$$X_C^P = \big\{A = (4,6),\ B = (8,2)\big\}.$$

2 шаг. Множество вершин прямоугольника

$$Y = \big\{y = (y_1,y_2)|\ 1 \le y_1 \le 8,\ 3 \le y_2 \le 6\big\}$$

есть

$$Y_C = \big\{c^{(1)} = (1,3), c^{(2)} = (1,6), c^{(3)} = (8,3), c^{(4)} = (8,6)\big\}.$$

Для каждого набора $c^{(i)} = \Big(c_1^{(i)}, c_2^{(i)}\Big) \in Y_C$ решаем оптимизационную задачу

$$\begin{gathered} f(x) = c_1^{(i)}x_1 + c_2^{(i)}x_2 \to \max, \\ \begin{cases} c^{(i)} = \Big(c_1^{(i)}, c_2^{(i)}\Big) \in Y_C,\ i \in \{1,2,3,4\}, \\ x = (x_1,x_2) \in X_C^P. \end{cases} \end{gathered} \tag{12}$$

Получаем множество оптимальных решений $x^*\big(c^{(i)}\big) = \Big(x_1^*(c^{(i)}), x_2^*(c^{(i)})\Big)$, $i \in \{1,2,3,4\}$, задач вида (12), именно

$$\big\{x^*(c^{(1)}) = (4,6),\ x^*(c^{(2)}) = (4,6),\ x^*(c^{(3)}) = (8,2),\ x^*(c^{(4)}) = (8,2)\big\}.$$

3 шаг. Пусть $x^*\big(c^{(i)}\big)$ оптимальные решения задачи (12) с заданным набором $c^{(i)} = \left(c_1^{(i)}, c_2^{(i)}\right) \in Y_C$. Составляем функции

$$\phi(x, c^{(i)}) = c_1^{(i)}\big(x_1^*(c^{(i)}) - x_1\big) + c_2^{(i)}\big(x_2^*(c^{(i)}) - x_2\big),\ i \in \{1, 2, 3, 4\}.$$

В результате имеем систему функций:

$$\begin{aligned} \phi(x, c^{(1)}) &= 22 - x_1 - 3x_2, \\ \phi(x, c^{(2)}) &= 40 - x_1 - 6x_2, \\ \phi(x, c^{(3)}) &= 70 - 8x_1 - 3x_2, \\ \phi(x, c^{(4)}) &= 76 - 8x_1 - 6x_2. \end{aligned} \tag{13}$$

4 шаг. Для каждого набора $c^{(i)} = \left(c_1^{(i)}, c_2^{(i)}\right) \in Y_C,\ i \in \{1, 2, 3, 4\}$, графически находим множество $X\big(c^{(i)}\big)$, которое определяется системой неравенств

$$\phi(x, c^{(i)}) \geq \phi(x, c^{(j)}),\ \forall i, j \in \{1, 2, 3, 4\},\ j \neq i.$$

Отметим, что система

$$\phi(x, c^{(1)}) \geq \phi(x, c^{(j)}),\ \forall j \in \{2, 3, 4\},$$

не имеет решений, то есть множество $X\big(c^{(1)}\big)$ пустое.

5 шаг. Получаем явный вид функции $\Phi_S[x] = \phi(x, c^{(i)}),\ c^{(i)} \in Y_C,\ i \in \{2, 3, 4\}$. В результате получаем представление множества X как объединение конечного числа подмножеств, именно,

$$X = \bigcup_{i \in \{2,3,4\}} \Big(X\big(c^{(i)}\big) \cap X\Big).$$

6 шаг. Решаем классические задачи линейного программирования вида (8), именно,

$$\begin{gathered} \Phi_S[x] = \sum_{k=1}^{2} c_k^{(i)}\Big(x_k^*\big(c^{(i)}\big) - x_k\Big) \to \min, \\ x \in \Big(X\big(c^{(i)}\big) \cap X\Big),\ i \in \{2, 3, 4\}. \end{gathered} \tag{14}$$

7 шаг. Окончательно, сравнивая решения задач вида (14), находим «глобальный» минимум функции $\Phi_S[x]$ на множестве X. Получаем, что план $x_S^* = (6, 4)$ является оптимальным решением задачи (9) согласно принципу Сэвиджа, то есть

$$\Phi_S[x_S^*] = \min_{x \in X} \max_{y \in Y} \phi(x, y) = \min_{x \in X} \Phi_S[x] = 10.$$

8 шаг. Учитывая (3), опишем функцию $R_S(x)$ на множестве X^P максимальных по Парето точек множества X:

$$R_S(x) = \begin{cases} 60 - 8x_1 - 3x_2,\ 4 \leq x_1 \leq 6,\ x_1 + x_2 = 10, \\ 30 - x_1 - 6x_2,\ 6 \leq x_1 \leq 8,\ x_1 + x_2 = 10. \end{cases}$$

Таким образом, на множестве X^P максимальных по Парето точек множества X задачи (9) функцию стратегического риска по Сэвиджу можно записать в виде

$$R_S(x) = \begin{cases} 30 - 5x_1,\ 4 \le x_1 \le 6,\ x_1 + x_2 = 10, \\ 5x_1 - 30,\ 6 \le x_1 \le 8,\ x_1 + x_2 = 10. \end{cases}$$

С учетом (11), функция стратегического риска по Вальду на множестве X^P примет вид

$$R_V(x) = 2x_1 - 8,\ 4 \le x_1 \le 8,\ x_1 + x_2 = 10.$$

Для отыскания U-оптимального по рискам и сожалениям решения задачи (9) найдем решение задачи минимизации функции $F(x) = R_V^2(x) + R_S^2(x)$ на множестве X^P, т. е.

$$F(x) = 29x_1^2 - 332x_1 + 964 \to \min,$$
$$\begin{cases} x_1 + x_2 = 10, \\ 4 \le x_1 \le 8. \end{cases}$$

Используя необходимые и достаточные условия существования экстремума, можно показать, что

$$x_U = arg \min_{x \in X} F(x) = \left(\frac{166}{29}, \frac{124}{29} \right).$$

Найденное решение является U-оптимальным по рискам и сожалениям в задаче (9) составления оптимального плана производства.

Заключение

В данной работе определяются понятия функции риска и функции сожаления для задачи линейного программирования при неопределенности. Каждое возможное решение ЛПР оценивает двумя критериями: стратегическим риском по Вальду и стратегическим сожалением по Сэвиджу. Первый критерий вычисляет возможные потери при отклонении от гарантированного по Вальду решения. Второй критерий учитывает возможность реализации «благоприятных» для ЛПР неопределенностей.

Формализовано U-оптимальное по рискам и сожалениям решение исходной линейной модели, которое соответствует поведению разумного игрока. Построен пошаговый алгоритм нахождения оптимального решения и проиллюстрирован на конкретном примере.

Список литературы

1. Бардин, А. Е., Житенева, Ю. Н. Риски и сожаления в игре с природой // Международный научный журнал "Спектральные и эволюционные задачи": Труды XXII Международной конференции "Крымская осенняя математическая школа—симпозиум". — Симферополь: ТНУ, 2012. — Т. 22. — С. 2–5.

   BARDIN, A. & ZHITENEVA, J. (2012) The risks and regrets in game with nature. *International Scientific Journal "Spectral and Evolution Problems: Proceedings of the Twenty Second Crimean Autumn Mathematical School-Symposium (KROMSH–2012)".* 22. p. 2–5.

2. Бардин, А. Е., Житенева, Ю. Н. Равновесия гарантий в одной бескоалиционной игре при неопределенности // Анализ, моделирование, управление, развитие экономических систем: сборник научных трудов VIII Международной школы—симпозиума АМУР-2014. — Симферополь: ТНУ, 2014. — С. 23–24.

   BARDIN, A. & ZHITENEVA, J. (2014) The equilibrium of guarantees in one non-cooperative game under uncertainty. *Analysis, Modeling, Management, Development of Economic Systems: Proceedings of the Eighth International School-Symposium AMUR–2014.* . p. 23–24.

3. Жуковский, В. И., Кудрявцев, К. Н., Смирнова, Л. В. Гарантированные решения конфликтов и их приложения. — М.: КРАСАНД, 2013. — 324 с.

   ZHUKOVSKIY, V. and CUDRYAVCEV, K. and SMIRNOVA, L. (2013) *Guaranteed conflict solutions and their applications.* Moscow: KRASAND.

УДК: 519.853.53:519.816.4

MSC2010: 91A10

СПОСОБ ГАРАНТИРОВАННОГО РАСПРЕДЕЛЕНИЯ ДЕНЕЖНЫХ СРЕДСТВ ПО ДВУМ ДЕПОЗИТАМ

© Ю. А. Бельских

Государственный гуманитарно-технологический университет
физико-математический факультет
кафедра математики и физики
ул. Зеленая, 22, Орехово-Зуево, 142611, Российская федерация
e-mail: *fozbelskih@rambler.ru*

© В. И. Жуковский

Московский государственный университет им. М.В. Ломоносова
факультет вычислительной математики и кибернетики
кафедра оптимального управления
Ленинские горы, МГУ, ВМК, 2 учебный корпус, Москва, ГСП-1, 119991, Российская Федерация
e-mail: *zhkvlad@yandex.ru*

© Л. В. Смирнова

Государственный гуманитарно-технологический университет
факультет информатики
кафедра информатики
ул. Зеленая, 22, Орехово-Зуево, 142611, Российская федерация
e-mail: *smirnovalidiya@rambler.ru*

Method of Guaranteed Distribution of Available Funds in two Deposits.

Belskih J. A., Zhukovskiy V. I., Smirnova L. V.

Abstract. Method of distribution of fixed sum of money in ruble and currency deposits is suggested. Lying at the junction of results of the theory of multicriteria problems under uncertainty and principle of Savage-Nichans minimax regret (decision making in onecriteria problem under uncertainty) this method allows to estimate the accrued (per year) amount of deposit distributed at the beginning of the year in two (indicated above) deposits. Moreover, concerning the exchange rate at the end of the year only bounds of changes are known.

Keywords: *multicriteria problem, uncertainty, Slater and Pareto maximum, criterion, alternative, minimax regret principle*

Введение

Рассматривается частный случай задачи распределения инвестиций по портфелям, за решение которой Гарри Максу Марковицу присуждена Нобелевская премия. Предлагаемая статья охватывает распределение денежных средств по двум депозитам: рублевому и валютному. Эта задача ранее решалась в однокритеральной подстановке (Жуковский В. И., Солдатова Н. Г., Молоствов В. С.) методами математического программирования. В настоящей статье задача рассматривается в двухкритериальной постановке и используется метод векторной оптимизации при неопределенности, предложенный В. И. Жуковским. В результате авторы сформулировали простой и конструктивный способ распределения вклада по двум депозитам.

1. Постановка задачи

Пусть у вкладчика в начале года на руках имеется некоторая сумма средств, которую он желает распределить между рублевым и валютным депозитами с тем, чтобы через год получить наращенную сумму вклада. Именно, целью вкладчика является такое распределение всей наличности в начале года, при которой *наращенная* сумма была бы возможно большей. Важными здесь являются два факта: во-первых, зная, как распределять каждый рубль, вкладчик будет знать, как распределять любую сумму средств; во-вторых, будущий курс валюты, естественно, неизвестен, но имеется диапазон его возможных изменений. В таком случае, будущий курс «выступает» как интервальная неопределенность, о которой известны лишь границы изменения, а какие-либо вероятностные характеристики отсутствуют (по тем или иным причинам).

Подходящей математической моделью представляется следующая двухкритериальная задача:

$$\Gamma = \langle Z = [0,1], Y = [a,b], \{f_i(z,y)\}_{i=1,2} \rangle,$$

где Z — множество альтернатив вкладчика z, определяющих, какую часть рубля z он вкладывает в рублевый депозит, тогда оставшуюся часть $1-z$ отправляет на валютный. Сами критерии, значения которых называют *исходами*, в задаче Γ имеют вид:

$$f_1(z,y) = z(1+r), \quad f_2(z,y) = \frac{1-z}{K}(1+d)y, \tag{1}$$

здесь r и d — процентные ставки по рублевому и валютному депозитам соответственно; K и y курс валюты в начале и конце года соответственно. Как уже упоминалось, о курсе валюты в конце года известен лишь диапазон изменений $y \in [a, b]$, где $0 < a < b$ — заданные постоянные. Итак, вкладчик в начале года вносит z-вую часть каждого рубля в рублевый, а остаток $1 - z$ — в валютный депозит, затем конвертирует $1-z$ в валюту $\frac{1-z}{K}$ и через год с помощью обратной конвертации $\frac{1-z}{K}(1+d)y$ валютный вклад по курсу $y \in [a, b]$ переводится в рубли. Статья [4] как раз и посвящена построению альтернативы z^*, при которой сумма вкладов становится через год наибольшей. Однако в [4] не учтены риски, возникающие при реализации такого z^*. Тем более отметим, что учет рисков является актуальной задачей. Подтверждает эту актуальность Нобелевская премия 1990 г., врученная Гарри Максу Марковицу [1] за новый подход к исследованию риска распределения инвестиций и диверсификации ожидаемых инвестиционных доходов. Идея такого подхода в настоящей статье распространяется на двухкритериальный случай задачи Г.

Перейдем к понятию риска (по Сэвиджу–Нихансу) [2, 3]. Для этого каждому критерию $f_i(z, y)$ $(i = 1, 2)$ из (1) поставим в соответствие его *функцию риска* (сожаления) по Сэвиджу–Нихансу [2, 3]

$$R_i(z, y) = \max_{u \in Z} f_i(u, y) - f_i(z, y) \quad (i = 1, 2), \tag{2}$$

и его *сильно гарантированный риск* (по Сэвиджу–Нихансу) при альтернативе $z \in Z$:

$$R_i[z] = \max_{y \in Y} R_i(z, y). \tag{3}$$

Аналогично для каждого критерия $f_i(z, y)$ $(i = 1, 2)$ из (1) определим его *сильную гарантию* [5]

$$f_i[z] = \min_{y \in Y} f_i(z, y) \quad (i = 1, 2). \tag{4}$$

Наконец, задаче Г с помощью (1)–(4) поставим в соответствие двухкритериальную «задачу гарантий»

$$\Gamma_2 = \langle Z, \{F_i[z] = f_i[z] - R_i[z]\}_{i=1,2} \rangle. \tag{5}$$

Приведем два понятия из теории многокритериальных задач:

– в Γ_2 альтернатива $z^S \in Z$ называется *максимальной по Слейтеру (слабо эффективной)* в двухкритериальной задаче (5), если при $\forall z \in Z$ несовместна система строгих неравенств $F_i[z] > F_i[z^S]$ $(i = 1, 2)$;

– альтернатива $z^P \in Z$ *называется максимальной по Парето (эффективной)* в (5), если при $\forall z \in Z$ несовместна система неравенств $F_i[z] \geq F_i[z^P]$ $(i = 1, 2)$, из которых хотя бы одно строгое.

Очевидно, что любая эффективная альтернатива является одновременно слабо эффективной, обратное, вообще говоря, не верно.

Определение. Тройку $(z^S; f[z^S], R[z^S])$ назовем *сильно гарантированным по исходам и рискам максимальным по Слейтеру решением исходной задачи* Г (СГИР-МС), если существуют

1) $f_i[z] = \min\limits_{y \in Y} f_i(z, y)$, $R_i[z] = \max\limits_{y \in Y} R_i(z, y)$ $(i = 1, 2)$,

2) альтернатива z^S максимальна по Слейтеру в задаче (5).

Напомним, что в приведенном выше определении

$$f[z] = (f_1[z], f_2[z]), \quad R[z] = (R_1[z], R_2[z]),$$
$$R_i[z] = \max_{y \in Y} R_i(z, y), \quad R_i(z, y) = \max_{u \in Z} f_i(u, y) - f_i(z, y) \quad (i = 1, 2).$$

Почему же в качестве «хорошего» решения Г предлагается СГИР-МС?

Во-первых, оно отвечает на вопрос о выборе действия: вкладчику предлагается следовать альтернативе z^S из тройки $(z^S; f[z^S], R[z^S])$.

Во-вторых, эта стратегия z^S «обеспечивает» при каждом $i = 1, 2$ исход $f_i(z^S, y)$ не меньший $f_i[z^S]$ с риском $R_i(z^S, y)$ не большим $R_i[z^S]$ при реализации любой неопределенности $y \in Y$ (то есть z^S «устанавливает» нижние границы для реализующихся при $z = z^S$ исходов и верхние границы для рисков, сопровождающих такую реализацию).

В-третьих, ситуация z^S реализует «самые большие» в векторном смысле (максимальные по Слейтеру) исходы и соответствующие им «минус» риски. Иначе говоря, не существует другой альтернативы $z \neq z^S$, при которой увеличивались бы все гарантии по исходам $f_i[z^S]$ и *одновременно* уменьшились бы все гарантии $R_i[z^S]$ по рискам.

Заметим, что объединение «во-вторых» и «в-третьих» является некоторым аналогом «действия» максиминной стратегии в однокритериальной задаче при неопределенности. Только внутренний минимум в максимине для Г заменяется на $\min\limits_{y \in Y} F_i(z, y)$ $(i = 1, 2)$, а внешний максимум — на максимум по Слейтеру. Здесь возможны *два направления* для дальнейших исследований. Первое из них: заменить оптимум по Слейтеру на оптимум по Парето, по Борвейну, по Джоффриону, конусную оптимальность и установить связи между такими разными решениями. Второе направление основывается на стремлении вкладчика к большим исходам, а взятые нами в приведенном определении гарантии «самые маленькие». Поэтому можно заменить скалярные минимумы (из внутреннего минимума в максимине) на векторные из числа выше перечисленных (тем самым увеличив для некоторых $i = 1, 2$ гарантии).

Замечание. Приведенное определение позволяет предложить следующий *конструктивный способ* построения СГИР-МС. Он сводится к четырем этапам.

Этап 1. По $f_i(z, y)$ найти $\max_{z \in Z} f_i(z, y) = \psi_i[y]$ и построить функцию риска (по Сэвиджу–Нихансу) для критерия $f_i(z, y)$, именно,

$$R_i(z, y) = \psi_i[y] - f_i(z, y) \quad (i = 1, 2).$$

Этап 2. Построить гарантии исходов $f_i[z] = \min_{y \in Y} f_i(z, y)$, рисков $R_i[z] = \max_{y \in Y} R_i(z, y)$ $(i = 1, 2)$ и затем критерии $F_i[z] = f_i[z] - R_i[z]$ $(i = 1, 2)$ задачи (5).

Этап 3. Для двухкритериальной вспомогательной задачи гарантий $\Gamma^a = \langle Z, \{F_i[z]\}_{i=1,2} \rangle$ вычислить максимальную по Слейтеру альтернативу z^S. Здесь будем применять или теорему Ю. Б. Гермейера [6, с. 66] или лемму Карлина [6, с. 71].

Этап 4. Найти двухкомпонентные вектора

$$f[z^S] = (f_1[z^S], f_2[z^S]), \quad R[z^S] = (R_1[z^S], R_2[z^S])$$

и с их помощью выписать искомую СГИР-МС $(z^S; f[z^S], R[z^S])$.

2. Явный вид гарантированного решения задачи диверсификации

Нижеследующее утверждение представляет собой центральный результат настоящей статьи.

Утверждение. *Явный вид СГИР-МС в исходной задаче* Γ *будет*

$$(z^S, f[z^S], R[z^S]) = (z^S; f_1[z^S], f_2[z^S]; R_1[z^S], R_2[z^S]) =$$

$$= \begin{cases} (0; 0, \dfrac{1+d}{K}a; 1+r, 0) \ \ \textit{при} \ \ \gamma \le a, \\ \left(\dfrac{\gamma+a}{2\gamma+a+b}; \dfrac{(\gamma+a)(1+r)}{2\gamma+a+b}, \dfrac{\gamma+b}{2\gamma+a+b}\dfrac{1+d}{K}a; \right. \\ \left. (1+r)\dfrac{\gamma+b}{2\gamma+a+b}, b\dfrac{1+d}{K}\dfrac{\gamma+a}{2\gamma+a+b}\right) \ \ \textit{при} \ \ a < \gamma < b, \\ \left(1; 1+r, 0; 0, \dfrac{1+d}{K}b\right) \ \ \textit{при} \ \ \gamma \ge b, \end{cases} \tag{6}$$

где $\gamma = \frac{1+r}{1+d}K$.

Доказательство. Доказательство разделим на две части. В первой установим справедливость (6) при условии $\gamma \le a$ и $\gamma \ge b$, во второй докажем (6) при условии $a < \gamma < b$.

Следуем четырем этапам из замечания.

Этап 1. С помощью (2) построим

$$R_1(z, y) = [\max_{z\in[0,1]} f_1(z)] - (1+r)z = (1+r) - z(1+r) = (1-z)(1+r),$$
$$R_2(z, y) = [\max_{z\in[0,1]} f_2(z, y)] - (1-z)\frac{1+d}{K}y = zy\frac{1+d}{K}.$$

Этап 2. На основе (4) и (3) найдем сильные гарантии исходов $f_i[z]$ и рисков $R_i[z]$ $(i = 1, 2)$:

$$f_1[z] = \min_{y\in[a,b]} z(1+r) = z(1+r); \ \ f_2[z] = \min_{y\in[a,b]} (1-z)\frac{1+d}{K}y = (1-z)\frac{1+d}{K}a;$$
$$R_1[z] = \max_{y\in[a,b]} R_1(z, y) = (1-z)(1+r); \ \ R_2[z] = \max_{y\in[a,b]} R_2(z, y) = z\frac{1+d}{K}b.$$

С их помощью определим критерии $F_i[z] = f_i[z] - R_i[z]$ $(i = 1, 2)$ в задаче (5):

$$F_1[z] = f_1[z] - R_1[z] = (2z-1)(1+r),$$
$$F_2[z] = f_2[z] - R_2[z] = \frac{1+d}{K}a - \frac{1+d}{K}(a+b)z.$$

Этап 3. Здесь выделим *два подслучая.* В первом ограничимся условиями $\gamma = \frac{1+r}{1+d}K \le a$ и $\gamma \ge b$. Во втором рассмотрим оставшийся вариант $a < \gamma < b$.

Первый подслучай $\gamma \leq a$ или $\gamma \geq b$. Для нахождения слабо эффективной альтернативы z^S в задаче (5) применим лемму Карлина [6, с. 71], согласно которой z^S, найденное из

$$\max_{z\in[0,1]}(F_1[z]+F_2[z]) = F_1[z^S]+F_2[z^S]$$

будет максимальна по Парето (и, следовательно, по Слейтеру) в (5). С учетом обозначения $\gamma = \frac{1+r}{1+d}K$ имеем

$$\max_{z\in[0,1]}(F_1[z]+F_2[z]) = \frac{1+d}{K}\left\{[2\gamma-(a+b)]z-\gamma+a\right\}.$$

Эта функция линейна по z, определена на отрезке $Z=[0,1]$ и поэтому максимум достигается на концах отрезка $[0,1]$ либо при $z=0$, либо при $z=1$; причем эти максимумы

$$\varphi(0) = F_1[0]+F_2[0] = \frac{1+d}{K}(a-\gamma) \;\text{ и }\; \varphi(1) = F_1[1]+F_2[1] = \frac{1+d}{K}(\gamma-b).$$

Имеют место две импликации

$$[a\geq\gamma] \Rightarrow [\varphi(0)>\varphi(1)],\;\; [\gamma\geq b] \Rightarrow [\varphi(0)<\varphi(1)].$$

В самом деле,

$$[a\geq\gamma] \Leftrightarrow \left[\frac{a+a}{2}\geq\gamma\right] \Rightarrow \left[\frac{a+b}{2}>\gamma\right] \Rightarrow [a-\gamma>\gamma-b] \Rightarrow$$

$$\Rightarrow \left[\varphi(0)=\frac{1+d}{K}(a-\gamma)>\varphi(1)=\frac{1+d}{K}(\gamma-b)\right].$$

Аналогично доказывается и вторая из приведенных импликаций.

Итак установили, что при $a\geq\gamma$ максимальной по Слейтеру в (5) является альтернатива $z^S=0$, а при $\gamma\geq b$ — слабо эффективной альтернативой будет $z^S=1$. Соответствующие сильные гарантии исходов и рисков сразу определяются из формул, полученных в процессе доказательства на этапе 2 и они приведены в (6) для случая $\gamma\leq a$ и $\gamma\geq b$.

Перейдем ко *второму подслучаю* $a<\gamma<b$. Здесь уже применим частный случай указанной выше теоремы Ю. Б. Гермейера [6, с. 66]: альтернатива z^S, найденная из

$$\max_{z\in[0,1]}\min_{i=1,2}F_i[z] = \min_{i=1,2}F_i[z^S], \tag{7}$$

максимальна по Слейтеру в задаче (5). Из (7) сразу получаем для построения z^S равенство $F_1[z^S] = F_2[z^S]$, откуда

$$z^S = \frac{\gamma + a}{2\gamma + a + b}, \;\; 1 - z^S = \frac{\gamma + b}{2\gamma + a + b},$$

и поэтому с помощью формул для сильных гарантий из этапа 2 устанавливаем справедливость (6) случай $a < \gamma < b$.

Заключение

Настоящая статья может быть ориентиром для вкладчика, желающего внести на годовой срок деньги в банк для их "наращивания". Для него естественен вопрос: какую часть суммы положить на рублевый депозит, и тогда оставшаяся часть попадет на валютный депозит?

В статье рекомендуется при ограничении $\frac{1+r}{1+d}K \leq a$ все деньги вложить в валютный депозит, и тогда через год каждый рубль ему принесет доход, не меньший $\frac{1+d}{K}a$ с нулевым риском (наверняка!).

Если же $\frac{1+r}{1+d}K \geq b$, то рекомендуем вкладчику вложиться в рублевый депозит, что принесет ему через год с нулевым риском доход от каждого вложенного рубля не меньше $1 + r$.

Наконец, при $a < \frac{1+r}{1+d}K < b$ вкладчик часть $\frac{\gamma+a}{2\gamma+a+b}$ от каждого рубля вносит в рублевый депозит, а остаток в валютный. Через год он получит гарантированно за каждый вложенный рубль наращенную сумму $\frac{(\gamma+a)(1+r)}{2\gamma+a+b}$ по рублевому вкладу и $\frac{\gamma+b}{2\gamma+a+b}\frac{1+d}{K}a$ — по валютному, причем первый из них с риском по Сэвиджу–Нихансу $(1+r)\frac{(\gamma+b)}{2\gamma+a+b}$ и второй — с риском в $b\frac{1+d}{K}\frac{\gamma+a}{2\gamma+a+b}$.

Здесь r и d процентные ставки по рублевому и валютному депозитам соответственно; K и у курсы валюты в начале и в конце года; верхний предел изменения курса валюты в конце года y постоянная $b > 0$, а нижний предел $a > 0$ вкладчику приходится прогнозировать.

Список литературы

1. MARKOWITZ, H. M. (1952) Portfolio Selection . *The Journal of Finance*. 7 (1). p. 77–91.

2. NICHANS, J. (1948) Zur Preisbildung bei ungewissen Erwartungen. *Schweizerische Zeitschrift fur Volkswirtschaft and Statistik*. 84 (5). p. 433–456.

   NICHANS, J. (1948) To the pricing under uncertain expectations. *Swiss journal for Economics and Statistics*. 84 (5). p. 433–456.

3. SAVAGE, L. J. (1951) The theory of statistical division. *Journal of the American Statistical Association.* 46 (253). p. 55–67.

4. ZHUKOVSKIY, V. I., MOLOSTVOV, V. S. and TOPCHISHVILI, A. L. (2014) Problem of multicurrency deposit diversification – three possible approaches to risk accounting. *International Journal of Operations and Quantitative Management.* 20 (1). p. 1–15.

5. Жуковский, В. И., Кудрявцев, К. Н. Уравновешивание конфликтов при неопределенности. II. Аналог максимина // Математические основы теории игр и приложения. — 2013. — 5. 2. — С. 3–45.

   ZHUKOVSKIY, V. I. and KUDRYAVTSEV, K. N. (2013) Solving conflicts under uncertainty. II. Analog of maximin. *Mathematical foundation of game theory and applications.* 5 (2). p. 3–45.

6. Подиновский, В. В., Ногин, В. Д. Парето–оптимальное решение многокритериальных задач. — М.: Физматлит, 2007. — 256 с.

   PODINOVSKIY, V. V. and NOGIN, V. D. (2007) *Pareto optimal solution of multicriteria problems.* Moscow: Fizmatlit.

УДК: 519.16

MSC2010: 90C27

ИСПОЛЬЗОВАНИЕ ДОПОЛНИТЕЛЬНОЙ ИНФОРМАЦИИ В ЗАДАЧАХ ДИСКРЕТНОЙ ОПТИМИЗАЦИИ ТИПА МНОГИХ КОММИВОЯЖЕРОВ

© М. С. Германчук

Крымский федеральный университет им. В. И. Вернадского
Таврическая академия
факультет математики и информатики
просп. Академика Вернадского, 4, Симферополь, 295007, Российская Федерация
e-mail: *m.german4uk@yandex.ru*

Information Exploration for Discrete Optimization Problems such as Multiple Traveling Salesman Problems.

Germanchuk M. S.

Abstract. The problem of analysis and synthesis of optimal flows like the flow of resources, information flow etc. have a vital importance for the real systems. As a rule, the mathematical models are given by different types of nets. These nets are graph structures with marked vertices and edges. They yield a number of Discrete Optimization (DO) problems most of which are NP-complete. If the information connected with a given DO problems is taken into account, then it is possible to design approximate and heuristic algorithms capable to manipulate complex large-scale data. The Multiple Traveling Salesman Problem (MTSP) and its more generalized version the Vehicle Routing Problem (VRP) are the most typical test problems. They are connected with the problems of the shortest path, Hamiltonian circuit, vertex-edge transformation, maximum cut etc. Real situations contain similar, rational, extreme statements of problems, which are important for both exact and approximate solutions. Approximate solutions are based on combinations of local heuristic algorithms. Extremal problems of analysis and synthesis on graphs should take into account knowledge, information, facts and precedents. A variety of problems is defined by the type of graphs simulating resource networks; the structure of graphs, their dimension, the possibility of decomposition; the nature of the objective functions and completeness of information concerning coefficients for the criteria; the ability to represent restrictions on the network as disjunctive normal forms. The problem is complicated by usage of additional information (knowledge) for restrictions and requires adaptation of existing algorithms.

In this work on the base of knowledge-oriented approach, we give an overview of existing results for discrete optimization problems such as TSP with restrictions, formulate new problems, and suggest algorithms to solve them. The preliminary classification is given. It is shown that the knowledge consideration about the network structure, salesmen objectives, and prohibitions leads to decomposition (cluster) algorithms. The further development is associated with the approach based on the usage of controlled intelligent agents (in particular, salesman agents). We consider the generalized multi-agent traveler salesmen problems taking into account a variety of knowledge, information, data needed for both intelligent agents control and local agent control, the algorithms of decomposition, clustering, analysis and synthesis of networks. The preliminary numerical calculations confirm a necessity in a wide range of algorithms involved in the optimal composition of met heuristics and filling control systems.

Keywords: *Multiple Traveling Salesman Problem, knowledge-oriented models, approximation algorithms*

Введение

Задачи моделирования сложных социальных, экологических, экономических систем требуют детального учета существующего в таких системах многообразия потоков. Информационные и ресурсные потоки связаны с различного типа сетями: транспортными, водо-, электро-, теплоснабжения; предоставления услуг; сбора, вывоза и переработки бытовых отходов; рекреационными и др. Такие сети представляются графовыми структурами, которым приписаны некоторые величины (длина, стоимость, время, интенсивность и т. п.). Задачи эффективного управления потоками в сетях носят как локальный, так и глобальный характер. Необходимость локального исследования непосредственно связана с естественной локальностью сетевых структур или является следствием сложности системы, необходимости ее кластеризации (декомпозиции). Практическая востребованость таких задач определяет актуальность исследования. Указанная проблематика наглядно проявляется в задачах маршрутизации, задачах типа многих коммивояжеров [6, 11, 16, 17, 18], имеющих широкие приложения.

Целью работы является разработка методов и алгоритмов приближенного решения знаниеориентированных сетевых задач типа многих коммивояжеров большой размерности и связанных с ними задач псевдобулевой оптимизации, служащих наполнением соответствующих интеллектуализированных систем управления или поддержки принятия решений.

При исследовании таких задач возникает многообразие математических моделей, определяемое необходимостью учета исходной (входной), априорной, поступаемой и извлекаемой информацией, ее полнотой, неопределенностью и т. п.

В работе [3] (и библиографии к ней) указывается на актуальность исследований сетевых транспортных задач, разнообразие их постановок. В [3, 4, 16, 17, 18], в зависимости от исходной информации и критериев, связанных с эффективностью структурных преобразований, рассматриваются задачи анализа и синтеза оптимальной транспортной сети. Проблемы разработки прикладных алгоритмов интеллектуализации обработки информации в сетевых задачах обозначены в [9, 13, 24].

Оптимизационные сетевые задачи приводят к моделям дискретной (ДО) и непрерывной оптимизации (НО) в условиях многокритериальности, большой размерности, с дополнительными ограничениями, связанными с требованиями к прохождению маршрутов, транспортным средствам (ТС) и т. п. Для разработки практического инструментария решения такого класса задач необходим учет информации, знаний, прецедентов; набор широкого класса точных, приближенных методов, метаэвристик и их композиций, учитывающих специфику задачи. Решающую роль здесь играет соотношение между локальными и глобальными методами и алгоритмами решения задач. Анализ, синтез, изменение структуры сети, ее кластеризация, выявление главных компонент требуют разработки итерационных оптимизационных процедур. Развитие сети [3] предполагает наличие соответствующих декомпозиционных методов ДО и методов реоптимизации [10], связанных с изменением параметров существующих элементов сети, с введением новых элементов (вершины, дуги), изменяющих структуру сети, ее топологию, характеристики компонент, связность. Характерной для такого подхода является задача для многих коммивояжеров.

В работе приведен обзор существующих результатов по задачам ДО типа коммивояжера с ограничениями и получены новые постановки задач, предложены алгоритмы их решения. Приведена предварительная классификация. Показано, что учет знаний о структуре сети, целях коммивояжеров, запретах приводит к декомпозиционным (кластерным) алгоритмам. Частично результаты докладывались на научно-практической конференции «Молодая наука» [25].

Задача коммивояжера (ЗК) (Travelling Salesman Problem – TSP) – задача нахождения оптимального маршрута обхода вершин графа с заданной матрицей попарных стоимостей перемещения между вершинами является классической NP-трудной задачей с широкими приложениями [5, 8, 9, 19, 20, 22, 23, 24, 25, 26, 27]. Рассматриваемый в работе класс задач является развитием классической ЗК. Обзор теоретических результатов, точных и приближенных алгоритмов решения ЗК содержится в работах [16, 17, 18] и библиографиях к ним. Приведем сетевую формулировку ЗК. Пусть $G = (V, U)$ – ориентированный (неориентированный) граф, где V $(|V| = n)$ – множество вершин, U – множество дуг (ребер). Для ориентированных графов приняты термины: дуга, путь, контур. Для неориентированных графов – ребро, цепь, цикл. Предполагается, что графы без петель и кратных дуг (ребер). Пусть C – матрица $n \times n$ действительных чисел, $c_{ij} \geq 0, (i, j) \in U$ длин дуг (весовых коэффициентов, стоимости). В ЗК требуется найти замкнутый маршрут (контур, цикл, т. е. начинающийся и заканчивающийся в одном городе) коммивояжера, проходящий через все города (вершины графа) по одному разу и имеющий минимальную стоимость (длину, если c_{ij} – расстояния).

Знаниеориентированная модификация ЗК зависит от исходной информации, дополнительных требований и условий.

В работе [14] предложен вариант решения задачи типа многих коммивояжеров с помощью построения максимальных разрезов. Декомпозиция задачи с помощью процедуры максимального разреза выделяет две компоненты, на которых решается ЗК. Тем самым определяется необходимое число коммивояжеров. Так как задача построения максимального разреза является NP-трудной, то в работе [14] проведен сравнительный анализ использования эвристических алгоритмов (ЭА) и показано, что лучшие результаты дает комбинация эвристик. Выбор наилучшей комбинации ЭА зависит от размерности задачи, полноты информации и возможности ее использования. Это предопределяет необходимость постоянного пополнения базы моделей задач типа коммивояжера, базы алгоритмов, базы знаний, в которой фиксируются знания о сети, коммивояжерах, ограничениях, критериях и целевых функциях в различных формах (ДНФ – дизъюнктивных нормальных формах, продукциях, фреймах и т. п.), а также прецеденты и результаты реальных и квазиреальных вычислительных экспериментах.

Задача построения максимальных разрезов, применяемая для решения задачи многих коммивояжеров, может рассматриваться как задача кластеризации или декомпозиции исходной задачи и наоборот. В зависимости от учета обязательных знаний о задаче могут применяться различные методы кластеризации (декомпозиции) [1], а не только методы построения максимальных разрезов.

Рассмотрим некоторые из ключевых моментов реализации программы исследований для указанного подхода на примере знаниеориентированной задачи двух коммивояжеров.

1. Обобщенный алгоритм приближенного решения ЗК большой размерности

Алгоритм приближенного решения ЗК большой размерности, допускающий широкие обобщения в зависимости от привлечения различной информации, можно представить в следующем виде:

1°. Задать сеть $< G, C >$ (граф и данные о распределении весовых коэффициентов C на дугах и информация о вершинах).

2°. Провести кластеризацию сети (множество вершин графа разбить на два $V = V_1 \cup V_2$, $V_1 \cap V_2 = \varnothing$, $G_i = (V_i, U_i)$, $i = 1, 2$).

3°. Построить кратчайшие маршруты обхода вершин V_1 и V_2 полученных сетей.

4°. Представить решение задачи для двух коммивояжеров.

5°. Используя 4, найти решение задачи для одного коммивояжера (если необходимо), как задачи обхода множеств.

Реализация каждого шага алгоритма существенно зависит от наличия доступной и обязательной информации (знаний о структуре, модели, целях и т. д.). Приведем некоторые варианты.

1. Задан граф $G = (V, U)$ (ориентированный, неориентированный) и матрица стоимостей (расстояний) C ($c_{ij} \geq 0$, $(i, j) \in U$, $i, j \in V$).

2. Граф задан частично и существует информация для восстановления составляющих, необходимых для решения ЗК.

3. Сеть меняется с течением времени:

– матрица C зависит от времени t, $c_{ij} = c_{ij}(t)$; множество вершин и дуг не меняется;

– множество вершин и дуг меняется с течением времени.

4. Сеть доступна локально:

– по мере продвижения агента-коммивояжера;

– из-за большой размерности сети для реализации решения задачи доступна только часть сети.

5. Сеть находится в развитии (синтез сети, старение сети, аварийное (катастрофическое) изменение сети).

6. Сеть задается с помощью процедуры распознавания сети (алгоритмически).

7. Задание сети зависит от целей, поставленных при решении исходной прикладной задачи, сводящейся к задаче типа m-коммивояжеров (ЗmК):

– цели определены полностью, непротиворечивы, иерархии целей соответствует иерархия критериев;

– цели определены не полностью и необходима процедура устранения неопределенности.

8. Цели и критерии разные у каждого коммивояжера или общие. Определена информация о критериях достижения целей.

9. Данные в сети полностью доступны каждому коммивояжеру или разная информация для каждого с непустым пересечением информационных множеств. Например, известна статистика о данных сети: распределение весов, длин дуг, стоимости, характеристики для вершин и т. п.

10. Учет знаний и прецедентов. Задание обязательных ограничений:

– по прохождению дуг (очередность) для каждого коммивояжера;

– по запретам на прохождение дуг каждым коммивояжером;

– по очередности посещения вершин;

– по сохранению оптимальных фрагментов маршрута, используемых ранее коммивояжером. Такие знания, как правило, можно формализовать в виде ДНФ ограничений. Приведем пример такой формализации.

Если дуга $(i, j) \in R \subset U$ принадлежит множеству дуг обязательных для включения в маршрут, то соответствующая переменная $x_{ij} = 1$. Для дуг, входящих в вершину i и выходящих из j, справедливы формулы: $\bigvee\limits_{(k,i)\in U} x_{ki} = 1$ и $\bigvee\limits_{(j,m)\in U} x_{jm} = 1$. Объединение приводит к ограничениям

$$\mathop{\&}\limits_{(i,j)\in R} \left(\left(\bigvee\limits_{(k,i)\in U} x_{ki} \right) x_{ij} \left(\bigvee\limits_{(j,m)\in U} x_{jm} \right) \right).$$

В работах [22, 23, 24] отражены направления, связанные с ограничениями, внутренними потерями и с условиями предшествования (задача курьера). Другие обобщения связаны с обходом конечной системы множеств и дополнительными затратами, связанными с выполнением работ. Для решения применяется оптимальный алгоритм на основе экономичной версии динамического программирования, в рамках которого и учитывается дополнительная информация. Заметим, что для такого класса задач метод ветвей и границ также позволяет учитывать локальную информацию (ограничения).

Возможные приложения связаны с транспортными перевозками, логистическими задачами, для которых характерны ограничения в виде условий предшествования. Транспортное средство по прибытию в пункт может принимать груз или корреспонденцию для доставки в другой пункт. Образуется пара: отправитель – получатель. Информация о таких парах формирует ограничения необходимые для исполнения, что усложняет выбор маршрута коммивояжера.

Примером учета знаний о вершинах в задаче двух коммивояжеров с множеством вершин V графа G является выделение множества $V_1 \subset V(|V_1| = k; |V| = n)$. Требуется построить маршрут коммивояжеров, обязательно проходящий через множество выделенных вершин V_1 точно один раз, а через остальные $(n - k)$ вершин – не более одного раза [6, 11].

Указанная постановка задачи допускает дальнейшие обобщения. Выделяются два множества вершин V_1 и V_2 :$V_1 \cup V_2 \subseteq V$. Нужно найти два гамильтоновых контура Γ_1 и Γ_2 такие, что агент вершины V_1 посещает по одному разу, а остальные не более одного, аналогично для другого агента-коммивояжера (вершины V_2 по одному разу, а остальные не более одного).

Видоизменение данной постановки заключается в выделении множества вершин V_3, которые коммивояжеры посещают не более одного раза: $V_1 \cup V_2 \cup V_3 \subseteq V$.

10.1. Известна прецедентная информация прохождения маршрутов (для всей или части сети) предпочтительная для лица, принимающего решение (ЛПР).

10.2. Есть дополнительная информация: априорная о решении, прецедентная, о необходимой точности решения и т. д.

11. Условия согласования по выбору: оптимальных маршрутов между коммивояжерами; обмену информацией; характеру взаимодействия (кооперативному, некооперированному, пошаговому и т. п.).

12. Известна или неизвестна информация о полноте знаний по всем сетевым компонентам.

Следующим вариантом является обобщенная ЗК; она состоит в построении полного кратчайшего контура, в котором на посещение вершины не накладывается условие точно один раз. Для обобщенной ЗК в такой формулировке возможны указанные выше обобщения.

В ЗК с условиями на дуги, выделяется фиксированное множество дуг $U_1 \subset U$. Требуется построить маршрут коммивояжера, обязательно проходящий через это множество дуг. Такие условия ставятся в задаче о сельском почтальоне, в задаче о сборе мусора.

В ЗК с несколькими весовыми матрицами кроме матрицы C задается еще одна или несколько других матриц весов дуг (ребер) [16]:

1) матрицы задают ограничения сверху и (или) снизу на вес маршрута коммивояжера и требуется найти решения ЗК с матрицей C, удовлетворяющие ограничениям на веса/маршрут (кроме длины маршрута, могут быть ограничения на время прохождения маршрута, общие затраты и т. п.);

2) матрицы задают дополнительные ограничения на части маршрута коммивояжера;

3) матрицы содержат критерии и ограничения.

2. Кластерная задача многих коммивояжеров

В практических задачах может быть неизвестно, какую необходимо иметь информацию о сети, что является следствием неудачных постановок задач, спецификаций, моделей и т. п. Роль адекватной, рациональной, целевой информации возрастает для многокомпонентных сетей.

2.1. **Варианты обхода множества вершин.** Например, в кластерной задаче многих коммивояжеров множество вершин V разбито на кластеры (компоненты) V_i, $i = \overline{1,k}$, $V_1 \cup V_2 \cup \ldots \cup V_k \subseteq V$. Возможны различные варианты обхода вершин:

1) j-й коммивояжер посещает все вершины кластера V_i подряд и только после этого переходит к межкластерному маршруту;

2) на маршрут в кластере V_i и между кластерами могут быть заданы условия предшествования и другие условия;

3) требуется посетить в каждом кластере только часть вершин (или одну, минимальный маршрут между V_i, $i = \overline{1,k}$);

4) находится представительная вершина кластера V_i, такая, что расстояние до непосещаемых вершин кластера не превышает заданной величины;

5) требуется определить кластеры, удовлетворяющие условиям 4);

6) между коммивояжерами возможна конкуренция и т. д. (игровые обобщения).

Выделение блоков в задачах ДО типа многих коммивояжеров можно интерпретировать как процесс кластеризации, который определяется спецификой задачи. В том случае, когда явно присутствуют кластеры (иерархическая инфраструктура города, региона и т. д.), можно использовать известные алгоритмы кластеризации [1]. Декомпозиция задачи с помощью процедуры максимального разреза выделяет две компоненты, на которых решается ЗК. Тем самым определяется необходимое число коммивояжеров. Если размерности полученных компонент велики, процедура повторяется. Для приближенного решения используются эвристики, как в задаче о максимальном разрезе, так и для решения ЗК на кластерах. Например, приближенное решение задачи о максимальном разрезе предполагает предварительное извлечение информации:

1) по статистике распределения весовых коэффициентов, расстояний (ранжирования);

2) по грубому распределению кластеров, качественной информации о компонентах, экспертной и прецедентной информации.

В работе [19] предлагается использовать декомпозиционный алгоритм, в котором с помощью какой-нибудь кластеризации множество вершин разбивается на некоторое число компонент (групп). Далее отыскивается рациональный обход этих множеств (компонент) и устанавливаются вершины входа и выхода для смежных компонент. Для каждой компоненты определяется кратчайший маршрут, соединяющий точки входа и выхода. Затем эти маршруты объединяются.

В задаче кластеризации учитывается удаленность вершин друг от друга. Предполагается что число кластеров и центры группирования (вершины) заданы. Процедура кластеризации проводится итерационно.

2.2. **Учет знаний о принадлежности выделенной дуги одному из коммивояжеров.** В задаче m-коммивояжеров принадлежность дуги (i, j) одному из них приводит к ситуациям:

а) дуга (i, j) принадлежит искомому пути одного из коммивояжеров, но (i, j) не принадлежит другим коммивояжерам;

б) одновременно принадлежит k-му и p-му коммивояжерам;

в) одновременно принадлежит всем коммивояжерам;

г) только одна из вершин — i и/или j принадлежит кому-то из других коммивояжеров.

2.3. **Алгоритм кластеризации задачи.** Алгоритм состоит из следующих шести шагов.

1°. В кластер первого коммивояжера включаем вершины i и j. Для второго коммивояжера выбираем вершину k максимально удаленную от вершин i и j.

2°. Находим максимальный разрез между вершинами i,j и вершиной k.

3°. Находим минимальный разрез, отделяющий вершины i,j от вершины k.

4°. Если цели коммивояжеров не противоположны, то кластеризацию проводим по максимальному разрезу. Тем самым в искомые маршруты не войдут дуги с максимальным суммарным весом.

5°. На кластерах строим пути (контуры) каждого из коммивояжеров.

6°. Если цели коммивояжеров противоположны, то второй коммивояжер может использовать вершины минимального разреза, тем самым, удлиняя путь первого коммивояжера. Здесь возможна различная тактика агентов-коммивояжеров в зависимости от целей и доступной информации (игры на графах).

Приведем применение одного из простых алгоритмов построения разрезов, который содержит следующие преобразования: найти максимальное значение (расстояние) $c_{ij} \geq 0$, если на дугу (i, j) не наложено ограничений, то вершину i включить в маршрут первого коммивояжера, а вершину j – в маршрут второго коммивояжера, иначе, взять следующую по рейтингу дугу (k, p) и повторить процедуру, а дугу (i, j) включить в маршрут одного из коммивояжеров и т. д. Параллельная работа трех агентов, два из которых коммивояжеры, а третий осуществляет декомпозицию ЗК для двух коммивояжеров позволяет упростить решения задачи.

При реализации такой схемы с обязательными ограничениями, дополнительной информацией на компонентах (множество меньшей размерности) и построении маршрута обхода множеств в качестве моделей ставится в соответствие ряд задач псевдобулевой условной оптимизации, в частности с дизъюнктивными ограничениями, для которых можно использовать результаты многокритериальной условной псевдобулевой оптимизации [12].

Реальные ситуации содержат близкие, рациональные, экстремальные (оптимизационные) постановки задач, для которых важны как точные, так и приближенные решения. Приближенные решения, как правило, базируются на комбинациях локальных эвристических алгоритмов. В экстремальных задачах анализа и синтеза на графах, как указывалось, необходимо учитывать знания, факты и прецеденты. Разнообразие задач диктуется классами графов, моделирующих ресурсные сети; структурой графов, их размерностью, возможностью декомпозиции; характером целевых функций и полнотой информации о коэффициентах критериев; возможностью представления знаний об ограничениях на сети в виде ДНФ. Использование дополнительной информации (знаний) по обязательным ограничениям усложняют задачу и требуют адаптации существующих алгоритмов.

Известные алгоритмы, реализующие построение кратчайшего пути, гамильтонового контура, максимального разреза, решения обобщенной задачи коммивояжера, вершинно-реберных преобразований, нахождения потоковых характеристик сети требуют модификаций, удовлетворяющих обязательным ограничениям, извлекаемым и возникающим в процессе решения задачи знаниям.

Исходя из сочетания локального и глобального характера знаний о задаче типа TSP в работах [6, 11], приведен подход по рациональному рассмотрению многокритериальных задач нескольких коммивояжеров как агентных в самоорганизующейся системе. Агентная задача многих коммивояжеров на изменяющейся сети естественно возникает, когда в рамках большой сетевой системы реализуется оптимальная сеть (синтез сети). На базе известных решений проводится реоптимизация относительно добавления вершин, дуг, внутренних условий (ДНФ ограничений), знаний, по количеству коммивояжеров-агентов. Устойчивость задач или оптимальных решений позволяет строить эффективные алгоритмы.

Заметим, что количество агентов-коммивояжеров связано с декомпозицией исходной задачи большой размерности, с выделением характерных кластеров (город, район и т. п.). С другой стороны решение задачи кластеризации может осуществляться с помощью агентов-коммивояжеров (задача распознавания конечных графов несколькими агентами).

Заключение

Базируясь на классической задаче типа многих коммивояжеров и применяя знаниеориентированный подход, удалось описать как известные, так и новые постановки задач дискретной оптимизации. Показано, что учет знаний, фактов, прецедентов может реализоваться в моделях условной псевдобулевой оптимизации. Обоснование алгоритмов может опираться на точные методы (ветвей и границ, динамического программирования). Эти методы позволяют учитывать различные условия предшествования, обязательного включения в оптимальный маршрут выделенного множества дуг и другие. Размерность сети и NP-сложность задач требуют эвристических подходов, основанных на различных алгоритмах декомпозиции и кластеризации. Такие алгоритмы предложены в работе. Также рассмотрены обобщенные многоагентные задачи типа коммивояжера, в которых учитываются разнообразные знания, данные, необходимые для интеллектуального управления агентами и локального управления, осуществляемого самим агентом, алгоритмы декомпозиции, кластеризации, анализа и синтеза сети. Предварительные численные расчеты подтверждают необходимость создания широкого комплекса алгоритмов, участвующих в оптимальной композиции метаэвристик и наполняющих системы управления.

Список литературы

1. Айвазян, С. А. Прикладная статистика и основы эконометрии. Учебник для вузов. — М.: ЮНИТИ, 1998. — 1022 с.

   AYVAZIAN, S. A. (1998) *Applied Statistics and Econometrics fundamentals. Textbook for high schools.* Moscow: UNITY.

2. Барханов, И. Ф. Об оптимальном приведении матрицы стоимостей / И. Ф. Барханов, В. Р. Фазылов // Ученые записки Казанского государственного университета. Физико-математические науки. — 2006. — Т. 148, кн. 2. — С. 18-22.

   BARKHANOV, I. F. and FAZYLOV, V. R. (2006) On the optimal present value matrix. *Scientific notes of the Kazan State University. Physics and mathematics.* T. 148, b. 2. p. 18-22.

3. Белоусова, Н. И. Оценка экономии от структуры в задачах анализа и синтеза транспортных сетей / Н. И. Белоусова, С. П. Бушанский, Е. М. Васильева // Анализ, моделирование, управление, развитие экономических систем: сборник научных трудов VIII Международной школы-симпозиума АМУР-2014, Севастополь, 12–21 сентября 2014 / под ред. доцента А. В. Сигала. — Симферополь: ТНУ им. В. И. Вернадского, 2014. — 360 с.

   BELOUSOVA, N. I., BUSHANSKIY, S. P. and VASIL'YEVA, Ye. M. (2014) *Evaluation of the structure of the economy in the analysis and synthesis of transport systems. Analysis, modeling, management, development of economic systems: Proceedings of the VIII International Symposium school-AMUR 2014.* Simferopol: TNU.

4. Белоусова, Н. И. Информационная технология синтеза сложных сетевых структур нестационарной российской экономики: модели, алгоритмы, программная реализация / Н. И. Белоусова, С. П. Бушанский, Е. М. Васильева, В. Н. Лившиц, Э. И. Позамантир // Аудит и финансовый анализ. — М.: ЗАО 1с: Компьютерный аудит. – Вып. 1. – 2008. — С. 50-88.

   BELOUSOVA, N. I., BUSHANSKIY, S. P., VASIL'YEVA, Ye. M., LIVSHITS, V. N. and POZAMANTIR, E. I. (2008) Information technology is the synthesis of complex network structures unsteady Russian economy: models, algorithms, software implementation. *Audit and financial analysis.* 1. p. 50-88.

5. Гаращенко, И. В. Метод решения гамильтоновой задачи коммивояжера / И. В. Гаращенко, А. В. Морозов, А. В. Панишев // Искусственный интеллект. — 2008. — Вып. 3. — С. 630-637.

   GARASHCHENKO, I. V., MOROZOV, A. V. and PANISHEV, A. V. (2008) A method for solving the traveling salesman problem Hamiltonian. *Artificial Intelligence.* 3. p. 630-637.

6. Германчук, М. С. Использование информации в задачах типа многих коммивояжеров / М. С. Германчук, М. Г. Козлова // XXV Крымская осенняя математическая школа-симпозиум по спектральным и эволюционным задачам (КРОМШ-2014). Тезисы докладов. – Симферополь: ТНУ, 2014. — № 1. — С. 65.

   GERMANCHUK, M. S. and KOZLOVA, M. G. (2014) Using the information in problems such as many business travelers. *KROMSH-2014.* 1. p. 65.

7. Дорн, Ю. В. Поиск неэффективных ребер в транспортных сетях // Труды МФТИ. — 2014. — Т. 6. – № 1. — С. 162-168.

   DORN, YU. V. (2014) Search inefficient edges in transport networks. *Proceedings of MIPT.* T. 6, b.1. p. 162-168.

8. Иванко, Е. Е. Минимаксная задача мультикоммивояжера с плавающим центром в исследовании эволюционной изменчивости // Управление в технических, эргатических, организационных и сетевых системах. Материалы конференции. Под ред. С. Н. Васильева, И. А. Каляева, Д. А. Новикова, Г. Г. Серебрякова. — 2012. — С. 1164-1167.

   IVANKO, Ye. Ye. (2012) Minimax multi Salesman problem with floating center in the study of evolutionary variability. *Management of Engineering, ergatic, organizational and network systems.* p. 1164-1167.

9. Иванко, Е. Е. Задача курьера как эвристика для решения задачи коммивояжера // Проблемы оптимизации и экономические приложения. Материалы VI Международной конференции. — 2015. — С. 124.

   IVANKO, Ye. Ye. (2015) Task carrier as a heuristic for solving the traveling salesman problem. *Problems of optimization and economic applications. Proceedings of the IV International Conference.* p. 124.

10. Иванко, Е. Е. Маршрутно-распределительные задачи: теория и приложения // Дисс. на соискание ученой степени доктора физ.-мат. наук. – Екатеринбург, 2015.

    IVANKO, Ye. Ye. (2015) Routing and distribution tasks: Theory and Applications. *The thesis for the degree of Doctor of Physical and Mathematical Sciences.* Ekaterinburg.

11. Козлова, М. Г. Обобщения задачи коммивояжера: знаниеориентированный подход / М. Г. Козлова, М. С. Германчук // Інформатика та системні науки (ІСН-2013): матеріали IV Всеукр. наук.-практ. конф., (м. Полтава, 21-23 берез. 2013р.) / за ред. Ємця О. О. – Полтава: ПУЕТ, 2013. – С. 147-150.

    KOZLOVA, M. G. and GERMANCHUK, M. S (2013) Generalized traveling salesman problem: knowledge-oriented approach. *Computer and System Sciences. Proceedings of the IV All-Ukrainian scientific-practical conference.* p. 147-150.

12. Козлова, М. Г. Многокритериальные модели принятия решений с линейными псевдобулевыми функциями и дизъюнктивными ограничениями / М. Г. Козлова // Искусственный интеллект. — 2000. — № 2. — С. 67-73.

    KOZLOVA, M. G. (2000) Multicriteria decision model with linear pseudo-disjunctive functions and limitations. *Artificial Intelligence.* 2. p. 67-73.

13. Козлова, М. Г. Прикладные алгоритмы интеллектуализации обработки информации в моделировании задач типа коммивояжера / М. Г. Козлова, М. С. Германчук // Анализ, моделирование, управление, развитие социально-экономических систем: сборник научных трудов IX Международной школы-симпозиума АМУР-2015, Севастополь, 12-21 сентября 2015 / Под ред. доцента А. В. Сигала. – Симферополь: КФУ имени В. И. Вернадского, 2015. – С. 161-164.

    KOZLOVA, M. G. and GERMANCHUK, M. S (2015) Applied algorithms intellectualization information processing tasks such as simulations of a Salesman. *Analysis, modeling, management, development of socio-economic systems: Proceedings of the IX International Symposium school.* p. 161-164.

14. Козлова, М. Г. Приближенное решение задачи о максимальном разрезе и ее применение / М. Г. Козлова, М. С. Германчук, Э. Д. Куртнебиев // Інформатика та системні науки (ІСН-2013): матеріали IV Всеукр. наук.-практ. конф., (м. Полтава, 21-23 берез. 2013р.) / за ред. Ємця О. О. – Полтава: ПУЕТ, 2013. – С. 150-153.

    KOZLOVA, M. G., GERMANCHUK, M. S and KURTNEBIYEV, E. D. (2013) An approximate solution of the problem of the maximum cross-section and its application. *Computer and System Sciences. Proceedings of the IV All-Ukrainian scientific-practical conference.* p. 150-153.

15. Левченко А. Ю., Морозов А. В., Панишев А. В. Быстрый алгоритм решения задачи о назначениях для нахождения нижней границы стоимости маршрута коммивояжера // Искусственный интеллект. — 2011. — № 4. — С. 406-416.

    LEVCHENKO, A. YU., MOROZOV A. V. and PANISHEV A. V. (2011) Fast algorithm for solving the assignment problem to find the lower limit value of the traveling salesman route. *Artificial Intelligence.* 4. p. 406-416.

16. Меламед, И. И. Задача коммивояжера. Вопросы теории / И. И. Меламед, С. И. Сергеев, И. Х. Сигал // Автоматика и телемеханика. — 1989. — № 9. — С. 3-34.

    MELAMED, I. I., SERGEYEV, S. I. and SIGAL, I. KH. (1989) The traveling salesman problem. Questions of theory. *Automation and Remote Control.* 9. p. 3-34.

17. Меламед, И. И. Задача коммивояжера. Точные алгоритмы / И. И. Меламед, С. И. Сергеев, И. Х. Сигал // Автоматика и телемеханика. — 1989. — № 10. — С. 3-29.

MELAMED, I. I., SERGEYEV, S. I. and SIGAL, I. KH. (1989) The traveling salesman problem. Exact algorithms. *Automation and Remote Control*. 10. p. 3-29.

18. Меламед, И. И. Задача коммивояжера. Приближенные алгоритмы / И. И. Меламед, С. И. Сергеев, И. Х. Сигал // Автоматика и телемеханика. — 1989. — № 11. — С. 3-26.

MELAMED, I. I., SERGEYEV, S. I. and SIGAL, I. KH. (1989) The traveling salesman problem. Approximate algorithms. *Automation and Remote Control*. 11. p. 3-26.

19. Серая О. В., Бачкир Л. В. Стохастическая задача коммивояжера // Вестник НТУ ХПИ. — 2006. — № 40. — С. 169-173.

SERAYA, O. V., and BACHKIR, L. V. (2006) Stochastic Traveling Salesman Problem. *Bulletin NTU KHPI*. 40. p. 169-173.

20. Сергеев, С. И. Гибридные системы управления и динамическая задача коммивояжера // Автоматика и телемеханика. — 2008. — № 1. — С. 45-54.

SERGEEV, S. I. (2008) The hybrid control system and dynamic traveling salesman problem. *Automation and Remote Control*. 1. p. 45-54.

21. Сигал И. Х., Иванова А. П. Введение в прикладное дискретное программирование: модели и вычислительные алгоритмы. — М.: Физматлит, 2007. — 304 с.

SIGAL, I. KH. and IVANOVA, A. P. (2007) *Introduction to Applied Discrete Programming: models and computational algorithms*. Moscow: Fizmatlit.

22. Ченцов А. А. Обобщенная модель курьера с дополнительными ограничениями / А. А. Ченцов, А. Г. Ченцов // Вестник Южно-Уральского государственного университета. Серия: Математическое моделирование и программирование. — 2016. — Т. 9. — № 1. — С. 46-58.

CHENTSOV, A. A. and CHENTSOV, A. G. (2016) The generalized model of the courier with additional restrictions. *Bulletin of South Ural State University. Series: Mathematical modeling and programming*. T. 9. № 1. p. 46–58.

23. Ченцов А. Г. Экстремальные задачи маршрутизации и распределения заданий: вопросы теории. — Москва-Ижевск: РХД, 2008. — 238 с.

CHENTSOV, A. G. (2008) *Extreme problems in routing and distribution of tasks: Theory*. Moscow-Izhevsk: RKHD.

24. Ченцов А. Г. Эффективный метод решения задачи обхода мегаполисов при ограничениях предшествования специального типа / А. Г. Ченцов, Д. М. Хачай // Современные проблемы математики и ее приложения. Труды 47-1 Международной Всероссийской школы-конференции. — 2016. — С. 191-199.

CHENTSOV, A. G. and KHACHAY, D. M. (2016) Effective method for solving the problem of megacities crawl under the constraints of precedence of a special type. *Contemporary Mathematics and its Applications. Proceedings 47-1 of the International School of the All-Russian conference.* p. 191-199.

25. Германчук М. С. Система интеллектуального управления в прикладных сетевых задачах / М. С. Германчук // Научно-практическая конференция «Молодая наука»: сборник трудов / под общей редакцией Н. Г. Гончаровой. — Симферополь: ИТ «АРИА». — 2015. — С. 46-48.

    GERMANCHUK, M. S (2015) Intelligent control systems in the application of network problems. *Scientific-practical conference «Young Science»*. p. 46–48.

26. GUTIN, G., PUNNEN, A. P. (2007) The travelling salesman problem and its variations. Dordrecht: Kluwer Academic Publishers.

27. VORONIN, A. N., KOZLOV, A. I. (1994) Multiobjective problems solution under uncertainty *Multiple criteria problems under uncertainty. Abstracts of The Third International Workshop*. p. 99.

Zhukovskiy V., Kirichenko M., Boldyrev M. & Smirnova L. V. 2016. Risks in a Multicriteria Problem under Uncertainty. *Taurida Journal of Computer Science Theory and Mathematics*, 1, pp. 7 – 22.

MSC2010: 90C29

Новизна подхода, предложенного в настоящей статье, состоит в том, что лицо, принимающее решение (ЛПР) в многокритериальной задаче при неопределенности стремится не только увеличить гарантированное значение каждого из своих критериев, но и одновременно уменьшить гарантированные риски, сопровождающие такое увеличение. Предлагаемое исследование выполнено на стыке теории многокритериальных задач и принципа минимаксного сожаления Сэвиджа–Ниханса. Из теории многокритериальных задач использовано понятие слабо эффективной оценки и сопровождающая теорема Гермейера Ю. Б., а из принципа минимаксного сожаления — оценка риска значением функции сожаления по Сэвиджу–Нихансу. Рассмотрение ограничено интервальными неопределенностями, о которых (ЛПР) известна лишь граница изменения, а какие-либо вероятностные характеристики отсутствуют. Введено новое понятие сильно гарантированного по исходам и рискам максимального по Слейтеру решения, установлено его существование при обычных для математического программирования ограничениях (непрерывность критериев, компактность множества стратегий и неопределенностей). В качестве приложения найден явный вид предложенного решения в задаче диверсификации вклада по рублевому и валютному депозитам.

Ключевые слова: *максимум по Парето, стратегия, неопределенность, векторная гарантия, риск по Сэвиджу, принцип минимаксного сожаления.*

Анафиев А. С., Блыщик В. Ф., Донской В. И. Устойчивость алгоритмов обучения классификации, основанных на модифицированной модели вычислении оценок / А. С. Анафиев, В. Ф. Блыщик, В. И. Донской // Таврический вестник информатики и математики. — 2016. — № 4 (33). — С. 23 – 47.

УДК: 519.95

В этой статье получен следующий теоретический результат: существует устойчивый алгоритм $\mathscr{A}$ обучения модифицированной модели ABO^*, гарантирующий её обучаемость в форме универсального эмпирического обобщения непосредственно по одной обучающей выборке путём минимизации эмпирического риска. Чтобы получить этот результат, была доказана LOO устойчивость алгоритма $\mathscr{A}$. Алгоритм $\mathscr{A}$ подробно описан в статье и является процедурой обучения с адаптацией, предполагающей варьирование только весов объектов обучающей выборки. Остальные параметры модели полагаются фиксированными. Этого оказалось достаточно, чтобы добиться требуемого результата. Предлагаемая модификация модели ABO минимальна: при вычислении оценок исключается только случай суммирования, когда объект "голосует за себя". Легко показать, что в случае, когда модифицированная модель ABO^* основана на использовании кратчайших элементарных логических отделителей (в частности — тупиковых тестов), универсальное эмпирическое обобщение будет также иметь место.

Ключевые слова: *модифицированная модель АВО*, устойчивость алгоритмов обучения, обучаемость.*

Бардин А. Е., Житенева Ю. Н. Линейная задача принятия решений с учетом рисков и сожалений / А. Е. Бардин, Ю. Н. Житенева // Таврический вестник информатики и математики. — 2016. — № 4 (33). — С. 48 – 58.

УДК: 519.87

В работе рассматривается задача линейного программирования при неопределенности, формализуется понятие U-оптимального по рискам и сожалениям решения данной задачи. Указан алгоритм построения оптимального решения. Реализация алгоритма показана на примере задачи составления оптимального плана производства

Ключевые слова: *задача линейного программирования, неопределенность, двухкритериальная задача, риск по Вальду, сожаление по Сэвиджу, минимум по Парето.*

Бельских Ю. А., Жуковский В. И., Смирнова Л. В. Способ гарантированного распределения денежных средств по двум депозитам / Ю. А. Бельских, В. И. Жуковский, Л. В. Смирнова // Таврический вестник информатики и математики. — 2016. — № 4 (33). — С. 59 – 67.

УДК: 519.853.53:519.816.4

Предлагается способ распределения фиксированной суммы средств по рублевому и валютному депозитам. Способ лежит на стыке результатов теории многокритериальных задач при неопределенности и принципа минимаксного сожаления Сэвиджа–Ниханса принятия решения в однокритериальной задаче при неопределенности. Этот

способ позволяет гарантированно оценить наращенную за год сумму вклада, распределенного в начале года по двум указанным депозитам. Причем о курсе валюты в конце года известны лишь границы изменения.

Ключевые слова: *многокритериальная задача, неопределенность, максимум по Слейтеру и Парето, критерий, альтернатива, принцип минимаксного сожаления.*

Германчук М. С. Использование дополнительной информации в задачах дискретной оптимизации типа многих коммивояжеров / М. С. Германчук // Таврический вестник информатики и математики. — 2016. — № 4 (33). — С. 68 – 82.

УДК: 519.16

Для реальных систем актуальной является проблема анализа и синтеза оптимальных потоков различной природы: ресурсных, информационных и других. В качестве математических моделей используются сети — графовые структуры, вершинам и дугам которых приписаны некоторые величины. Возникает многообразие классов задач дискретной оптимизации (ДО), как правило, NP-трудных. Естественный учет информации, связанной с данными задачами ДО, позволяет строить алгоритмы (приближенные, эвристические), пригодные для сложных задач большой размерности. Характерными и тестовыми являются задачи маршрутизации, задачи типа многих коммивояжеров. С ними связаны задачи построения кратчайшего пути, гамильтонова контура, вершинно-реберных преобразований, максимального разреза и другие. В реальных ситуациях возникают экстремальные постановки задач, для которых важны как точные, так и приближенные решения. Приближенные решения базируются на комбинациях локальных и эвристических алгоритмов. В экстремальных задачах анализа и синтеза на графах необходимо учитывать знания, факты, прецеденты и другую релевантную информацию. Разнообразие задач диктуется классами графов, моделирующих ресурсные сети; структурой графов, их размерностью, возможностью декомпозиции; характером целевых функций и полнотой информации о коэффициентах критериев; возможностью представления знаний об ограничениях на сети в виде дизъюнктивных нормальных форм. Использование дополнительной информации (знаний) по обязательным ограничениям усложняют задачу и требуют адаптации существующих алгоритмов.

В работе, используя знаниеориентированный подход, приведен обзор существующих результатов по задачам ДО типа коммивояжера с ограничениями и получены

новые постановки задач, предложены алгоритмы их решения. Приведена предварительная классификация. Показано, что учет знаний о структуре сети, целях коммивояжеров, запретах приводит к декомпозиционным (кластерным) алгоритмам. Дальнейшее развитие связывается с подходом, основанным на использовании управляемых интеллектуальных агентов (в частности, агентов-коммивояжеров). Рассмотрены обобщенные многоагентные задачи типа коммивояжера, в которых для интеллектуального управления учитывается разнообразная информация, необходимая для интеллектуального управления агентами, и локального управления, осуществляемого самим агентом, алгоритмы декомпозиции, кластеризации, анализа и синтеза сети. Предварительные численные расчеты подтверждают необходимость создания широкого комплекса алгоритмов, участвующих в оптимальной композиции метаэвристик и наполняющих системы управления.

Ключевые слова: *задачи типа многих коммивояжеров, знаниеориентированные модели, приближенные алгоритмы.*

СПИСОК АВТОРОВ НОМЕРА

Анафиев Айдер Сератович — к. ф.-м. н., доцент кафедры информатики факультета математики и информатики Таврической академии Крымского федерального университета им. В. И. Вернадского, г. Симферополь, Российская Федерация
e-mail: aydera@mail.ru

Бардин Александр Евгеньевич — к. ф.-м. н., доцент кафедры информатики факультета информатики Государственного гуманитарно-технологического университета, г. Орехово-Зуево, Российская Федерация
e-mail: inth2006@rambler.ru

Бельских Юлия Анатольевна — к. ф.-м. н., доцент кафедры физики и математики физико-математического факультета Государственного гуманитарно-технологического университета, г. Орехово-Зуево, Российская Федерация
e-mail: fozbelskih@rambler.ru

Блыщик Владимир Федорович — к. ф.-м. н., доцент кафедры информатики факультета математики и информатики Таврической академии Крымского федерального университета им. В. И. Вернадского, г. Симферополь, Российская Федерация
e-mail: veb@land.ru

Болдырев Михаил Владиславович — студент кафедры оптимального управления факультета вычислительной математики и кибернетики Московского Государственного Университета им. М. В. Ломоносова, г. Москва, Российская Федерация
e-mail: m_ boldyrev@list.ru

Германчук Мария Сергеевна — ассистент кафедры информатики факультета математики и информатики Таврической академии Крымского федерального университета им. В. И. Вернадского, г. Симферополь, Российская Федерация
e-mail: m.german4uk@yandex.ru

Донской Владимир Иосифович	д. ф.-м. н., профессор, заведующий кафедрой информатики факультета математики и информатики Таврической академии Крымского федерального университета им. В. И. Вернадского, г. Симферополь, Российская Федерация *e-mail: donskoy@tnu.crimea.ua*
Житенева Юлия Николаевна	к. ф.-м. н., доцент кафедры информатики факультета информатики Государственного гуманитарно-технологического университета, г. Орехово-Зуево, Российская Федерация *e-mail: ulya_zhiteneva@mail.ru*
Жуковский Владислав Иосифович	д. ф.-м. н., профессор кафедры оптимального управления факультета вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова, г. Москва, Российская Федерация *e-mail: zhkvlad@yandex.ru*
Кириченко Михаил Михайлович	студент кафедры оптимального управления факультета вычислительной математики и кибернетики Московского Государственного Университета им. М. В. Ломоносова, г. Москва, Российская Федерация *e-mail: moyomylo11@gmail.com*
Смирнова Лидия Викторовна	к. ф.-м. н., доцент кафедры информатики факультета информатики Государственного гуманитарно-технологического университета, г. Орехово-Зуево, Российская Федерация *e-mail: smirnovalidiya@rambler.ru*

Подписано к печати 30.11.2016. Формат 38x30/2. Бумага тип ОП. Объем 9,4 п. л. Тираж 50 экз.
Заказ № НП/45. Свободная цена.
Отпечатано в издательском отделе КФУ имени В. И. Вернадского
295007, г. Симферополь, пр. Академика Вернадского, 4