# ASSESSING THE EFFECTIVENESS OF SCENARIOS OBTAINED BY SIMULATION MODELING

I. V. Chernov

Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia

✉ ichernov@gmail.com

**Abstract.** This paper considers the challenges of increasing the effectiveness of a scenario approach to planning and management processes in the development of socio-economic systems under uncertainty. Methods are proposed to assess the effectiveness of scenario management based on encoding and comparing scenarios with a target scenario. For comparison, scenarios are formally represented, using a special algorithm, as a sequence of unique event identifiers. The assessment methods involve the event similarity of scenarios, the similarity of scenarios considering operations over a sequence of events, and the similarity of events based on the qualitative behavior parameters of significant factors. An important feature of scenario management is that the goal is formulated as the final state of the system and, moreover, as a sequence of events constituting the target scenario. The effectiveness of management options and the related scenarios is assessed using a target scenario (benchmark): the results of scenario modeling and the corresponding management options are compared with the benchmark. Scenario analysis is intended to support the management process of an event leading to the goal; therefore, it is reasonable to use as a benchmark the most rational scenario in terms of several characteristics with target values for the event being conducted. Scenario events are encoded using the scenario-event approach developed previously. This approach makes the transition from a real vector space (reflecting changes in the factor values of a scenario simulation model) to an event space. Practical examples are provided: the novel approaches to assessing the effectiveness of scenarios are used in several applied problems of ensuring the sustainable and safe development of complex socio-economic systems.

**Keywords**: scenario model, scenario, management, effectiveness, factor, event, simulation modeling, assessment, measure of similarity, measure of connectivity.

## INTRODUCTION

A clear mathematical formalization is required when addressing applied problems of planning and managing the development of complex systems and providing intelligent support for the preparation and implementation of managerial decisions. Such a formalization is intended to construct systems for studying various aspects of situation development based on the generation of alternative development scenarios or the behavior of various types of complex objects.

In this paper, the scenario analysis technology methodologically relies on the mathematical model of directed graphs, which is an extension of the classical graph model.

We define a functional graph ($f$-graph) as a tuple $<(X, E), V, W>$ [1], where:

1. $G = (X, E)$ is a directed graph (digraph);

2. $V: X \to \mathbb{R}$ or $V = \{V_x, x \in X\}$ is the set of vertex parameters, where each vertex $x \in X$ is associated with a certain real parameter $V_x$; $\mathbb{R}$ is the set of real numbers;

3. $W: E \times V \to \mathbb{R}$ or $W = \{W_e: V \to \mathbb{R}, e \in E\}$ is the set of arc weights. Each arc $e_{ij} \in E$, connecting vertices $v_i$ and $v_j$, is associated with a functional relationship.

Particular cases of functions reflecting the arc weights of graph models are given in Table 1.

The factor set $V$ of a scenario model is formed based on the extended vector space $Z$ of endogenous and exogenous variables [1].

Let events be defined through changes in the behavioral dynamics of a set of significant factors for situation analysis and management:

$$V^{(\mathrm{sig})} \subseteq V = \{v_n\},\ n = 1, 2, \ldots, N,$$

where $N$ is the number of factors in the scenario model.

According to [2], when presenting simulation results, it is possible to automatically analyze them on given time intervals by identifying two sets (local minima and local maxima) within the intervals, building the corresponding regression lines, and determining their slopes ($\propto^{\min}, \propto^{\max}$) relative to the simulation step axis (Fig. 1).

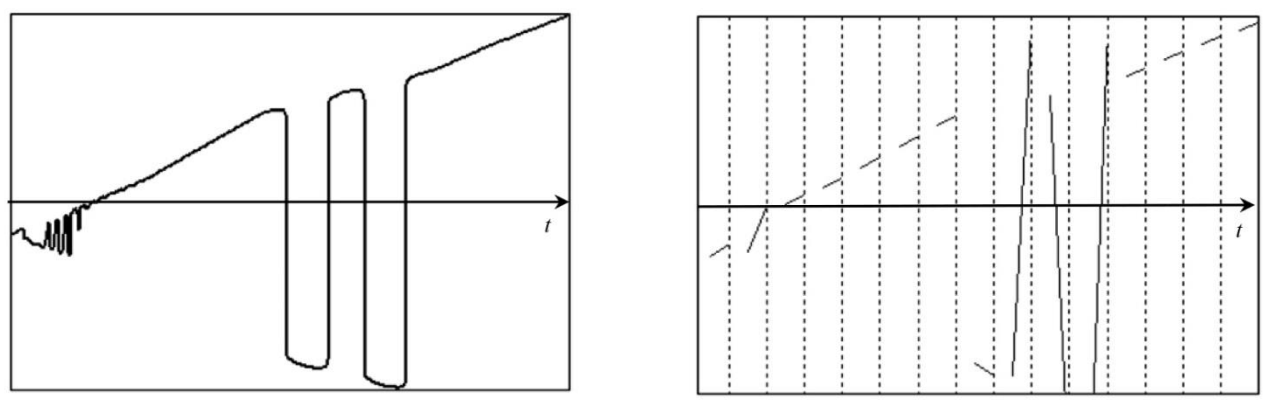
Also, such an analysis leads to a transition from a real vector space of scenario model's factor changes to an *f*-scenario space $FSc(v_i, t)$, which is characterized by six types of factor states (dynamics); see Table 2.

Table 2 presents the basic set of factor dynamics. First, this set can be expanded by combinations of the basic dynamics; second, it is possible to form new dynamics based on the characteristics of the factor change curves obtained by the analysis (e.g., based on the values of their slopes).

*Table 1*

**Particular cases of graph models**

| Name | Formalized representation |
|---|---|
| Signed digraph | $F\ (v_i, v_j, e_{ij}) = \begin{cases} +1 \text{ if an increase (decrease) of } v_i \\ \quad \text{causes an increase (decrease) of } v_j \\ -1 \text{ if an increase (decrease) of } v_i \\ \quad \text{causes a decrease (increase) of } v_j. \end{cases}$ |
| Weighted graph | $F(v_i, v_j, e_{ij}) = \begin{cases} +W_{i,j} \text{ if an increase (decrease) of } v_i \\ \quad \text{causes an increase (decrease) of } v_j \\ -W_{i,j} \text{ if an increase (decrease) of } v_i \\ \quad \text{causes a decrease (increase) of } v_j, \end{cases}$ where $W_{ij}$ is the weight of arc $e_{ij}$. |
| Functional graph | $F(v_i, v_j, e_{ij}) = f_{ij}(v_i, v_j, \ldots)$ |




(a) (b)

**Fig. 1. An example of building regression lines for the simulation results of a single factor in the scenario model:** (a) simulation results and (b) regression lines.

*Table 2*

**The identifiers of factor dynamics**

| The value of *f*-scenario $FSc(v_i, t)$ | Designation | The dynamics type of factor *i* | Default name |
|---|---|---|---|
| 0 | | Not defined (not calculated) | Not defined |
| 1 | | Increase | Increases |
| 2 | | Decrease | Decreases |
| 3 | | Constant value | Constant |
| 4 | | Fluctuations with bounded amplitude | Stable state |
| 5 | | Divergent fluctuations | Unstable |
| 6 | | Convergent fluctuations | Stabilizes |

Thus, a transition is performed to a discrete state space in which factor changes are described. In other words, for each significant factor, we have a finite set of discrete states that can change at discrete time instants, determined by the simulation step and the rules for calculating *f*-scenarios.

## 1. ENCODING OF SCENARIOS

In view of the above considerations, an expert-significant event at the *i*th step $t_i$ is a function of the behavioral dynamics of the significant factors for the formation of this event:

$$E^{(\mathrm{sig})}(t_i) = f\left(FSc(v_n, t_i)\right),\ v_n \in V^{(\mathrm{sig})},$$
$$n = 1, 2, \ldots, N,\ i = 1, 2, 3, \ldots$$

The simplest form of this function is the "concatenation" of all values $FSc(v_n, t_i)$. For example, if $V^{(\mathrm{sig})} = \{v_3, v_5, v_9, v_{12}\}$, $FSc(v_3, t_i) = 1$, $FSc(v_5, t_i) = 3$, $FSc(v_9, t_i) = 2$, and $FSc(v_{12}, t_i) = 1$ with $N = 12$ factors in the scenario model, then there are two concatenation variants:

- $E^{(\mathrm{sig})}(t_i) = \text{“1321”}$, the short representation, if the composition of significant factors $V^{(\mathrm{sig})}$ does not change over time;
- $E^{(\mathrm{sig})}(t_i) = \text{“}{-}{-}1{-}3{-}{-}{-}2{-}{-}1\text{”}$, the full representation, if the composition of significant factors may change: $V^{(\mathrm{sig})}(t)$.

Consider the case where a new event is formed when changing the dynamics type of any factor from the set of significant ones. Thus, we write the operator $\mathrm{create}\left(Ev_e\left(t_n, t_n + P\right)\right)$, creating a new event on the interval between the simulation steps $t_n$ and $t_n + P$:

$$\mathrm{create}\left(Ev_e\left(t_n, t_n + P\right)\right),\ e = 1,$$

$$\{v_i\} \in V^{(\mathrm{sig})},\ i = 1, 2, \ldots, N,$$

$$FSc(v_i, P) = 0,\ i = \overline{1, n},\ \text{if}\ t_n = 0.$$

$$t_n = t_n + P,\ \mathrm{create}\left(Ev_e\left(t_n, t_n + P\right)\right),\ e = e + 1,$$

$$\text{if}\ \exists FSc\left(v_i, t_n + P\right) \neq FSc\left(v_i, t_n\right).$$

Let a scenario be defined only by a sequence of events, which is the most common case.

Then the problem of determining the similarity of the scenario obtained by the scenario investigation to a target scenario is formulated as follows.

**Initial conditions.** On the set of scenarios, there are two given scenarios (or fragments of two scenarios), i.e., a target scenario $\mathfrak{R}^0$ and another one from the set of scenarios obtained under different conditions, $\mathfrak{R}^j \in \breve{\mathfrak{R}} = \{\mathfrak{R}^j\}$, $j = 1, 2, \ldots, N^{\mathfrak{R}}$, on the same time interval of duration $t_r$ at a time instant $t$, as finite sequences of ordered events $E^{(\mathrm{sig})j}$, formed by a rule $C^{(\mathrm{sig})}$ in the same vector space:

– a target scenario of the form

$$\mathfrak{R}^0(t) = \mathfrak{R}\left\{E^{(\mathrm{sig})0}(t_i) = f\left(Z^{(\mathrm{tar})}(t)\right), (t - t_r) < t_i \le t\right\};$$

– the *j*th scenario from the set of obtained scenarios of the form

$$\mathfrak{R}^j(t)=\mathfrak{R}\left\{E^{(\mathrm{sig})j}(t_i)=f(Z(t)),(t-t_r)<t_i\le t\right\}.$$

Here, $Z(t)$ is the state in the vector space at a time instant $t$, and $Z^{(\mathrm{tar})}(t)\subset Z(t)$ is a given target subspace.

**Problem:** it is required to define a measure of similarity of scenarios as that of event sequences.

We denote by

$$E^{(\mathrm{sig})j}(1)=E^{(\mathrm{sig})j}(t-t_r+1)$$

the first event of the scenario $\mathfrak{R}^j(t)$ obtained at step $t$ on the interval under consideration.

Thus, we consider the similarity of finite sequences of scenario events $\mathfrak{R}^j(t)$ $=\left\{E^{(\mathrm{sig})j}(1), E^{(\mathrm{sig})j}(2),\ldots, E^{(\mathrm{sig})j}(K)\right\}$ as objects representing ordered sets of $K=t_r$ elements $E^{(\mathrm{sig})j}(k)$ of a certain event set $\breve{E}^{(\mathrm{sig})}$, where $k$ is the event position on the scenario horizon $t_r$ (the relative simulation step number), $k=\overline{1, K}$.

Let $\breve{\mathfrak{R}}^{+}=\mathfrak{R}^0\cup\breve{\mathfrak{R}}$ be the set of all (target and obtained) scenarios; the target one is at the beginning of this set and is considered first. For the event-by-event comparison of scenarios, each event from the set $\breve{E}^{(\mathrm{sig})}$ is assigned an integer $TEv^{(\mathrm{sig})}$ (the event type), which serves as its unique identifier to distinguish this event from other noncoincident events in the set $\breve{E}^{(\mathrm{sig})}$. In this way, it is possible to form an array *AEv* containing the unique events $E^{(\mathrm{sig})}(t_i)$ of all scenarios from the set $\mathfrak{R}$ detected during the identification process. The number of an array element corresponds to the identifier (type) $TEv^{(\mathrm{sig})}$.

Unique events from the target and obtained scenarios are placed in the array as they are detected, and their identifiers (the numbers of these events in the array) are added to the representation $TEv^{(\mathrm{sig})j}$ of the current scenario. If an event has already been encountered (i.e., matches one of the events already present in the array $AEv$), it is assigned the identifier of that existing event in the array. In other words, a generalized algorithm (Fig. 2) is implemented step by step.

In Fig. 2 and subsequent figures, for brevity, the notation $E^{(\mathrm{sig})j}(t_i)\notin AEv$ means that the event $E^{(\mathrm{sig})j}(t_i)$ is not found among the elements of the array $AEv$.

Thus, scenarios can be represented as sequences of event types:

$$T\mathfrak{R}^j(k)=\left\{TEv^{(\mathrm{sig})j}(k)\right\},\ k=1, 2,\ldots, K,\quad j=0, 1, 2,\ldots, N^{\mathfrak{R}}. \tag{1}$$

Here, $\forall TEv^{(\mathrm{sig})j}(k)\in\{1,2,\ldots,|AEv|\}$, indicating a transition to the discrete states of events in scenarios. The number of types determines the number of events (scenario states) of the system modeled. Hence, scenario analysis can be performed on a finite set of discrete states of the system in discrete time (at simulation steps).

According to [2], there are six basic types of calculated *f*-scenarios and type 0 (no calculation), i.e., $FSc(v_i, t)\in\{0, 1, 2, 3, 4, 5, 6\}$. Therefore, the number $NEv^{(\mathrm{sig})}$ of possible expert-significant events for $Nv^{(\mathrm{sig})}$ event-generating factors is given by

$$NEv^{(\mathrm{sig})}=7^{Nv^{(\mathrm{sig})}}.$$

(This formula determines the number of all possible permutations of seven elements with repetitions.)

Note that for each factor $v_i$, each type of *f*-scenario can be assigned a name or a detailed description, generally also depending on $t$: $Name\_FSc(v_i, h, t)$, where $h=0, 1, 2, 3, 4, 5, 6$. Table 2 provides the default names for the dynamics types of factors.

The number of dynamics types can be expanded. For example, type 1 can be divided into two types: strict increase ($\propto^{\max}=\propto^{\min}$) or increase within a corridor ($\propto^{\max}>\propto^{\min}$), with possible fluctuations. In this case, the linguistic variable reflecting the dynamics of a factor $v_i$ is defined as

$$\left\{X(v_i),T_F(v_i),\theta^{\max}(v_i),\theta^{\min}(v_i),G_F(v_i),M_F(v_i)\right\},$$

where:

$X(v_i)$ is "the dynamics of a factor $v_i$";

$\theta^{\max}(v_i)=[-90, 90]$, $\theta^{\min}(v_i)=[-90, 90]$;

$T_F(v_i)$ is the term set of the variable $X$ (the names of linguistic values); by default, $\forall v_i$: $T_F(v_i)$ = {"increases", "decreases", "constant", "stable state", "unstable", "stabilizes"} with a membership function defined by Table 2;

$G_F(v_i)$ is a syntactic rule that generates new terms using quantifiers (e.g., "strongly", "weakly", "more or less", etc.); $M_F(v_i)$ is a procedure that maps each new term from $G_F(v_i)$ to an element of a fuzzy set $t_F\in T_F(v_i)$.

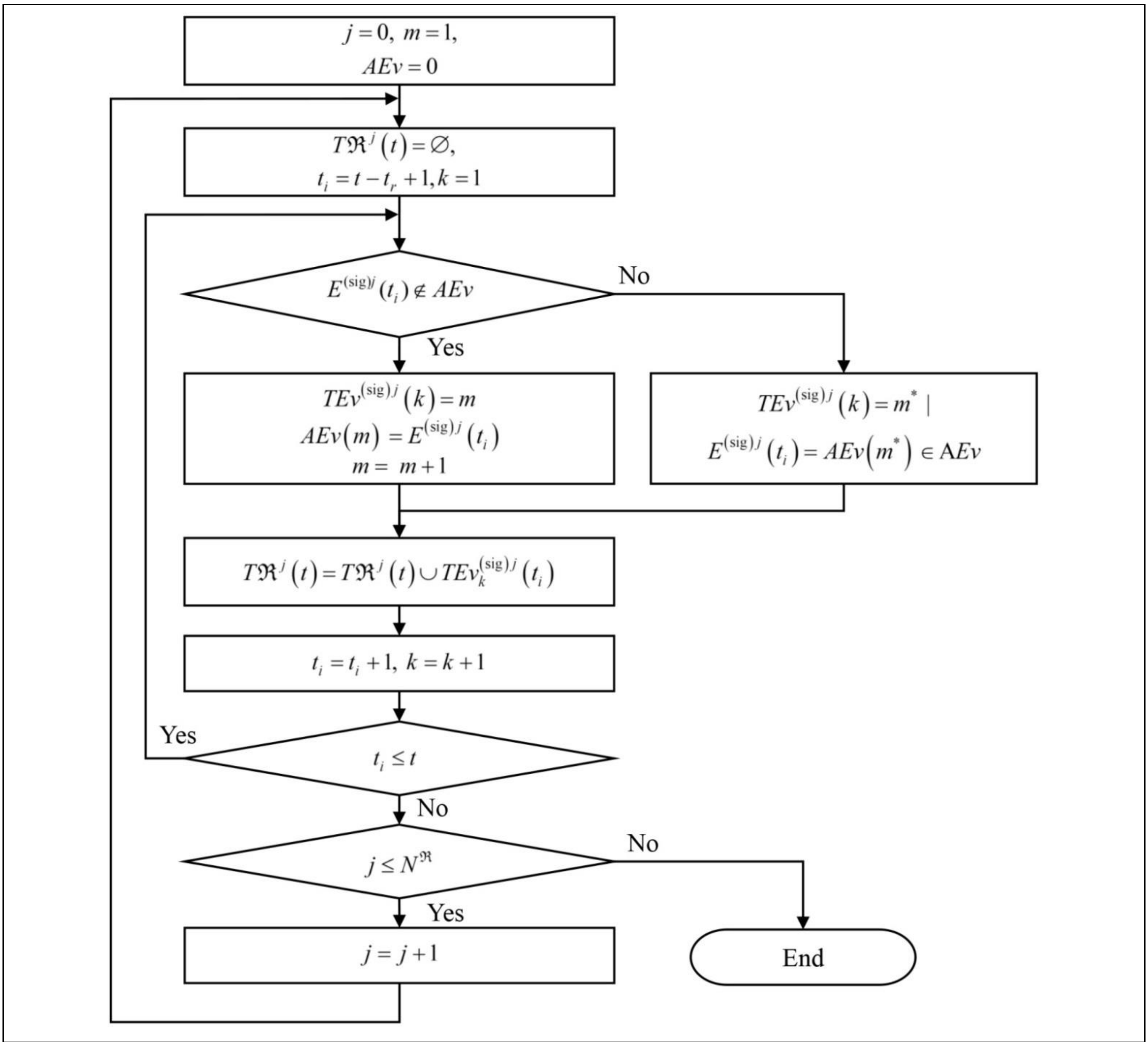



**Fig. 2. The algorithm for representing the target scenario as a sequence of unique event types.**

For example, a new term for the quantifier "strong + $t_F$" will have a membership function of the form $|\theta^{\max}(v_i)|>60,\ \ |\theta^{\min}(v_i)|>60$, etc.

The initial real vector space, defined by the values of factors $V: X\to\mathbb{R}$ or $V=\{V_x,\ x\in X\}$, is supplemented and, in most cases, replaced by the *f*-scenario space of discrete factor states $V^{\text{sc}}: X^{\text{sc}}\to FSc$ or $V^{\text{sc}}=\{V^{\text{sc}}_{x^{\text{sc}}},\ x^{\text{sc}}\in X^{\text{sc}}\}$. This transition improves the dynamic flexibility of scenario models (using results in structural element functions) and allows for the automatic generation of scenarios in text form. The results of scenario modeling can be transferred to other intelligent decision support systems. According to scenario modeling practice, in most cases, the number of distinct event types obtained during the modeling of a range of scenarios does not exceed a two-digit number.

Let each event be defined by a function of the *f*-scenarios specified in the target scenario and automatically identified in the simulation results $FSc(v_i,\ t_n+P)$, where $v_i\in V^{(\text{sig})}$. Then the number of event types is determined by the range (value set) of this function. In other words, the set of factors used to calculate *f*-scenarios forms a feature vector of the event set [3].

In the scenario modeling complex [4], event types are also denoted by natural numbers. Thus, a scenario can be described by a word in the alphabet of event types and a separator symbol (e.g., “,”), which is a subset of the set of natural numbers $\mathbb{N}^*$ (excluding zero[1]) and the separator symbol. Then the similarity of two scenarios, excluding type 0 (an undefined event), is assessed using the representation of scenarios as words in the alphabet

$$C^{\mathfrak{R}} = \{\text{“,”}, 1, 2, 3, 4, 5, \ldots\} = \mathbb{N}^* \cup \text{“,”}.$$

Consider an example of scenario modeling based on a model of geopolitical disagreements in the Arctic; its structure—the set of factors and relations between them—is shown in Fig. 3. This model was implemented in *Polyus*, a special program-analytical complex (PAC) for simulation modeling with the Russian-language interface. The description and scenario analysis of the model under different conditions and management options (control actions) were presented in [4].

Figure 4 shows an example of scenario modeling results, where events are determined based on the *f*-scenarios of eight selected factors.

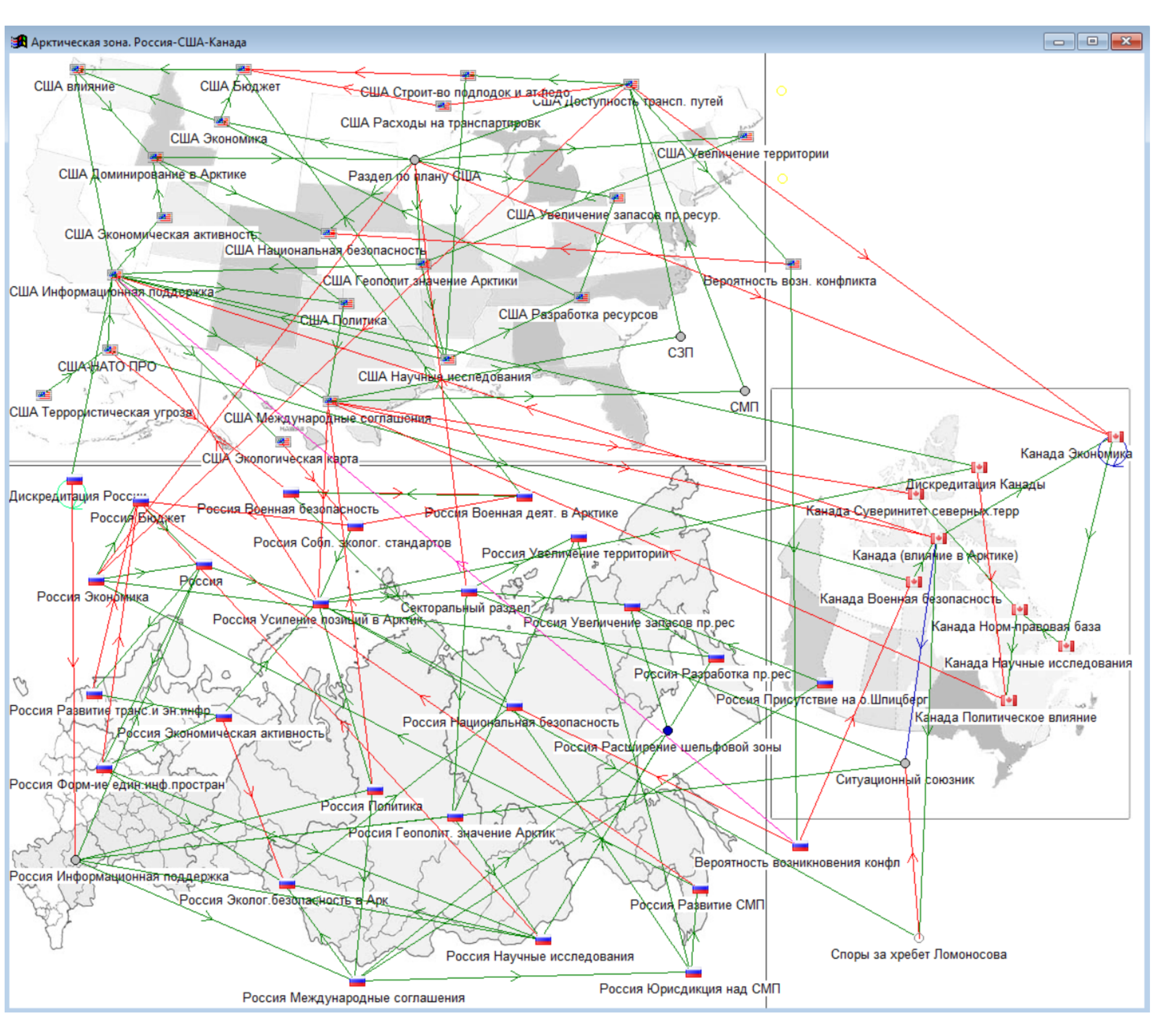



**Fig. 3. An example of the scenario model structure.**

[1] Hereinafter, we utilize the notation of *ISO 31-11* (1992), *ISO 80000-2* (2009), and *GOST R 54521-2011* (2011) standards.


**Fig. 4. An example of scenario modeling results.**

The state scale displays numerical designations for event types, in this case from 1 to 8. In the example under consideration, the event types (see Table 2) are distributed across the simulation steps as illustrated in Fig. 5. A change in the dynamics of at least one key factor marks the completion of the current event and the beginning of a new one. Figure 5 provides a fragment of the event-by-event text of the scenario automatically generated by PAC *Polyus*; the factors and their dynamics leading to the formation of a new event are in uppercase letters, and the factors with unchanged *f*-scenarios (the same as for the previous event) are in lowercase letters.

Thus, in this example with $t = t_r = 72$,

$$TEv^{(\text{sig})}(t) = \begin{cases} 1 \text{ for } t = \overline{1,5} \\ 2 \text{ for } t = \overline{6,10} \\ 3 \text{ for } t = \overline{11,15} \\ 2 \text{ for } t = \overline{16,20} \\ 4 \text{ for } t = \overline{21,40} \\ 5 \text{ for } t = \overline{41,60} \\ 6 \text{ for } t = \overline{61,67} \\ 7 \text{ for } t = \overline{68,72}. \end{cases}$$

In the simplest case of assessment, it suffices to represent a scenario as a sequence of different event types only, discarding the time intervals with identical types. That is, adjacent events in such a scenario representation differ in their type.

Thus, a scenario can be written as a word over the alphabet $C^{\mathfrak{R}}$:

- when considering all events (with possible recurrence) on continuous time intervals: 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 2, 2, 2, 2, 2, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 6, 6, 6, 6, 6, 6, 6, 7, 7, 7, 7, 7; or in the short notation 1(5)2(5)4(20)5(20)6(7)7(5), where the number of steps for each event type is indicated in parentheses (alternatively, as a two-dimensional array);
- when considering only the sequences of events that differ from their neighbors: 1, 2, 3, 2, 4, 5, 6, 7.

Note that, in the general case, events of the same type may be repeated in a scenario. For instance, in the current example, event type 2 appears twice in the scenario.

The measure of similarity [5] of scenarios is defined by a function $Mh\left(\breve{E}^{(\text{sig})} \times \breve{E}^{(\text{sig})}\right)$ with the following properties: $Mh \ge 0$, and for an event-by-event comparison of scenarios $\mathfrak{R}^i$ and $\mathfrak{R}^j$, $Mh\left(E_n^{(\text{sig})i}, E_n^{(\text{sig})j}\right) = Mh\left(E_n^{(\text{sig})j}, E_n^{(\text{sig})i}\right)$.

| Время | | Описание события | $TEv^{(\text{з}}$ |
|---|---|---|---|
| от | до | | |
| 1 | 5 | Динамика на рассчитывалась | 1 |
| 6 | 10 | "США ДОМИНИРОВАНИЕ В АРКТИКЕ" - УМЕНЬШАЕТСЯ<br>"США Национальная безопасность" - постоянно<br>"ДИСКРЕДИТАЦИЯ РОССИИ" - УМЕНЬШАЕТСЯ<br>"РОССИЯ ВОЕННАЯ ДЕЯТ. В АРКТИКЕ" - УМЕНЬШАЕТСЯ<br>"РОССИЯ УСИЛЕНИЕ ПОЗИЦИЙ В АРКТИКЕ» - УМЕНЬШАЕТСЯ<br>"РОССИЯ НАЦИОНАЛЬНАЯ БЕЗОПАСНОСТЬ" - УМЕНЬШАЕТСЯ<br>"РОССИЯ ИНФОРМАЦИОННАЯ ПОДДЕРЖКА" - РАСТЁТ<br>"ВЕРОЯТНОСТЬ ВОЗНИКНОВЕНИЯ КОНФЛ" - СТАБИЛИЗИРУЕТСЯ | 2 |
| 11 | 15 | "США ДОМИНИРОВАНИЕ В АРКТИКЕ" - НЕСТАБИЛЬНО<br>"США НАЦИОНАЛЬНАЯ БЕЗОПАСНОСТЬ" - НЕСТАБИЛЬНО<br>"ДИСКРЕДИТАЦИЯ РОССИИ" - ПРОЦЕСС НЕСТАБИЛЬНЫЙ<br>"РОССИЯ ВОЕННАЯ ДЕЯТ. В АРКТИКЕ" - НЕСТАБИЛЬНО<br>"РОССИЯ УСИЛЕНИЕ ПОЗИЦИЙ В АРКТИКЕ» - НЕСТАБИЛЬНО<br>"Россия Национальная безопасность" - уменьшается<br>"РОССИЯ ИНФОРМАЦИОННАЯ ПОДДЕРЖКА" - НЕСТАБИЛЬНО<br>"ВЕРОЯТНОСТЬ ВОЗНИКНОВЕНИЯ КОНФЛ" - НЕСТАБИЛЬНО | 3 |
| 16 | 20 | "США ДОМИНИРОВАНИЕ В АРКТИКЕ" - УМЕНЬШАЕТСЯ<br>" США НАЦИОНАЛЬНАЯ БЕЗОПАСНОСТЬ " - ПОСТОЯННО<br>"ДИСКРЕДИТАЦИЯ РОССИИ" - УМЕНЬШАЕТСЯ<br>"РОССИЯ ВОЕННАЯ ДЕЯТ. В АРКТИКЕ" - УМЕНЬШАЕТСЯ<br>"РОССИЯ УСИЛЕНИЕ ПОЗИЦИЙ В АРКТИКЕ» - УМЕНЬШАЕТСЯ<br>"Россия национальная безопасность" - уменьшается"<br>РОССИЯ ИНФОРМАЦИОННАЯ ПОДДЕРЖКА" - РАСТЁТ<br>"ВЕРОЯТНОСТЬ ВОЗНИКНОВЕНИЯ КОНФЛ" - СТАБИЛИЗИРУЕТСЯ | 2 |
| 21 | 40 | "США ДОМИНИРОВАНИЕ В АРКТИКЕ" - НЕСТАБИЛЬНО<br>"США НАЦИОНАЛЬНАЯ БЕЗОПАСНОСТЬ" - НЕСТАБИЛЬНО<br>"ДИСКРЕДИТАЦИЯ РОССИИ" - ПРОЦЕСС НЕСТАБИЛЬНЫЙ<br>"РОССИЯ ВОЕННАЯ ДЕЯТ. В АРКТИКЕ" - НЕСТАБИЛЬНО<br>"РОССИЯ УСИЛЕНИЕ ПОЗИЦИЙ В АРКТИКЕ» - РАСТЁТ<br>"РОССИЯ НАЦИОНАЛЬНАЯ БЕЗОПАСНОСТЬ" – РАСТЁТ<br>"РОССИЯ ИНФОРМАЦИОННАЯ ПОДДЕРЖКА" - РАСТЁТ<br>"ВЕРОЯТНОСТЬ ВОЗНИКНОВЕНИЯ КОНФЛ" - УМЕНЬШАЕТСЯ | 4 |
| 41 | 60 | "США ДОМИНИРОВАНИЕ В АРКТИКЕ" - УМЕНЬШАЕТСЯ<br>"США НАЦИОНАЛЬНАЯ БЕЗОПАСНОСТЬ" - УМЕНЬШАЕТСЯ<br>"Дискредитация России" - процесс нестабильный<br>"Россия Военная деят. в Арктике" - нестабильно<br>"Россия Усиление позиций В АРКТИКЕ» - растёт<br>"Россия Национальная безопасность" - растёт<br>"Россия Информационная поддержка" - растёт<br>"Вероятность возникновения конфл" - уменьшается | 5 |
| 61 | 67 | "США Доминирование в Арктике" - уменьшается<br>"США Национальная безопасность" - уменьшается<br>"ДИСКРЕДИТАЦИЯ РОССИИ" - УМЕНЬШАЕТСЯ<br>"Россия Военная деят. в Арктике" - нестабильно<br>"Россия Усиление позиций В АРКТИКЕ» - растёт<br>"Россия Национальная безопасность" - растёт<br>"Россия Информационная поддержка" - растёт<br>"Вероятность возникновения конфл" - уменьшается | 6 |
| 68 | 72 | "США Доминирование в Арктике" - уменьшается<br>"США Национальная безопасность" - уменьшается<br>"Дискредитация России" - уменьшается<br>"РОССИЯ ВОЕННАЯ ДЕЯТ. В АРКТИКЕ" - УМЕНЬШАЕТСЯ<br>"Россия Усиление позиций В АРКТИКЕ» - растёт<br>"Россия Национальная безопасность" - растёт<br>"Россия Информационная поддержка" - растёт<br>"Вероятность возникновения конфл" - уменьшается | 7 |

**Fig. 5. Event types and simulation steps.**

By a common assumption, $0 \le Mh \le 1$, where 0 and 1 mean complete dissimilarity and complete similarity, respectively.

The problem of determining an effective scenario is formulated as follows. Let $\mathfrak{R}^0$ be a given target scenario and $\breve{\mathfrak{R}} = \{\mathfrak{R}^j\}$, $j = 1, 2, \ldots, N^{\mathfrak{R}}$, be a given set (spectrum) of scenarios; in this set, it is required to find an effective scenario, i.e., the one most similar to the target scenario.

We will search for an effective scenario by comparing pairwise each of the obtained scenarios with the target one, using their measures of similarity.

Below are examples of various assessments of similarity that can be used to solve this problem.

## 2. SCENARIO ASSESSMENT BASED ON THE SIMILARITY OF EVENT SEQUENCES

Let us assess two scenarios based on the similarity of events $k = \overline{1, K}$ at each discrete time instant $t_i$, where $(t - t_r) \le t_i \le t$.

The similarity of two scenarios $\mathfrak{R}^0$ and $\mathfrak{R}^j$ $(\mathfrak{R}^j \in \breve{\mathfrak{R}},\ j = 1, 2, \ldots, N^{\mathfrak{R}})$ at a time instant $t$ can be measured, e.g., using a modification of the Hamming distance $d_H(\mathfrak{R}^0, \mathfrak{R}^j)$ [6]:

$$d_H(\mathfrak{R}^0, \mathfrak{R}^j, t) = \sum_{k=t-t_r+1}^{t} \left| TEv^{(\mathrm{sig})0}(k) - TEv^{(\mathrm{sig})j}(k) \right|. \tag{2}$$

In this paper, events are compared by their type. For brevity, $\delta_H$ refers to $\delta_H(\mathfrak{R}^0, \mathfrak{R}^j)$, without explicit specification of the arguments.

Essentially, we use the similarity of two words (scenarios) defined on an alphabet $C = \{TEv^{(\mathrm{sig})0}(k)\} \cap \{TEv^{(\mathrm{sig})j}(k)\}$ (the space of event types). To compare a set of scenarios with a target, the alphabet may be expanded, incorporating new events $C = \{TEv^{(\mathrm{sig})0}\} \cap \{TEv^{(\mathrm{sig})1}\} \cap \ldots \cap \{TEv^{(\mathrm{sig})j}\}$ at each union step. With this formulation, it is possible to apply the existing string comparison algorithms [7, 8].

Thus, at a given time instant $t$, a pair $\left(T\mathfrak{R}(t),\, d_H(t)\right)$, consisting of a set $T\mathfrak{R}(t)\in T\mathfrak{R}^i(t)$ of scenarios (1) represented as event types (1) and their pairwise comparison function $d_H\left(\mathfrak{R}^i,\, \mathfrak{R}^j,\, t\right)$: $T\mathfrak{R}(t)\times T\mathfrak{R}(t)$, which maps the Cartesian square of the set to the set of real numbers, is called a scenario metric space. Indeed, the following conditions (axioms) of a metric space are satisfied:

1. $d_H\left(\mathfrak{R}^i, \mathfrak{R}^j,\, t\right)=0 \Leftrightarrow \mathfrak{R}^i=\mathfrak{R}^j$ (identity);

2. $d_H\left(\mathfrak{R}^i,\, \mathfrak{R}^j,\, t\right)>0$ (nonnegativity);

3. $d_H\left(\mathfrak{R}^i,\, \mathfrak{R}^j,\, t\right)=d_H\left(\mathfrak{R}^i,\, \mathfrak{R}^j,\, t\right)$ (symmetry);

4. $d_H\left(\mathfrak{R}^i,\mathfrak{R}^j,t\right)=d_H\left(\mathfrak{R}^i,\mathfrak{R}^m,t\right)+d_H\left(\mathfrak{R}^m,\mathfrak{R}^j,t\right)$ (triangle).

In this case:

- $T\mathfrak{R}(t)$ is the domain, or the carrier, of the scenario metric space;
- $d_H(t)$ is the metric, or distance function;
- the scenarios $\mathfrak{R}^i$ are points of the scenario metric space.

In the above example, two scenarios are all the more similar the fewer events they differ by on the time axis, limited by the scenario depth. In the two extreme cases, the event sequences of two scenarios $\mathfrak{R}^0$ and $\mathfrak{R}^j$ may either coincide completely or differ completely. These considerations lead to the possibility of using known measures. A simple example of one such measure is the Hamming distance, which is defined in this work through the number of different events for the scenarios obtained by simulation on the same time interval. In this case, the measure of similarity $\delta_H$ is defined as

$$\delta_H\left(\mathfrak{R}^0,\mathfrak{R}^j\right)=1-\frac{d_H\left(\mathfrak{R}^0,\mathfrak{R}^j,K\right)}{K}. \qquad (3)$$

However, for the distance $d_H\left(\mathfrak{R}^0,\, \mathfrak{R}^j,\, k\right)$ (2), the distinction between different pairs of event types will vary significantly: for example, the distinction between increase and decrease will be $|1-2|=1$, while the distinction between increase and fluctuations with bounded amplitude will be $|1-4|=3$. To assess the mismatch between scenarios $\mathfrak{R}^0$ and $\mathfrak{R}^j$, it is reasonable to analyze only the cases of noncoincident event types. For this purpose, the numerical operation $\left|TEv^{(\mathrm{sig})0}(k)-TEv_k^{(\mathrm{sig})j}(k)\right|$ in formula (2) should naturally be replaced by the binary comparison function:

$$\theta\left(TEv^{(\mathrm{sig})0}(k),\, TEv^{(\mathrm{sig})j}(k)\right) =\begin{cases}0 \text{ if } TEv^{(\mathrm{sig})0}(k)=TEv^{(\mathrm{sig})j}(k)\\ 1 \text{ if } TEv^{(\mathrm{sig})0}(k)\neq TEv^{(\mathrm{sig})j}(k).\end{cases} \qquad (4)$$

The expression (4) considers extreme binary cases: if the events coincide, then $\theta=0$; if they differ, then $\theta=1$. In other words, $\theta\in\{0, 1\}$. However, when assessing the similarity of scenarios in practice, it is sometimes necessary to use fuzzy similarity (i.e., $\theta\in[0, 1]$), or an assessment of dissimilarity. Thus, an event pair may have greater similarity than another one, even despite the absence of complete similarity with the corresponding target event. In the general case, it is reasonable to introduce the following weight matrix of event (event type) distinctions at a simulation step $t$:

$$c^{\theta}(t)=\begin{bmatrix} 0 & c^{\theta}_{1,2}(t) & c^{\theta}_{1,3}(t) & \dots & c^{\theta}_{1,M-1}(t) & c^{\theta}_{1,M}(t)\\ c^{\theta}_{2,1}(t) & 0 & c^{\theta}_{2,3}(t) & \dots & c^{\theta}_{2,M-1}(t) & c^{\theta}_{2,M}(t)\\ \dots & \dots & \dots & \dots & \dots & \dots\\ c^{\theta}_{M-1,1}(t) & c^{\theta}_{M-1,2}(t) & c^{\theta}_{M-1,3}(t) & \dots & 0 & c^{\theta}_{M-1,M}(t)\\ c^{\theta}_{M,1}(t) & c^{\theta}_{M,2}(t) & c^{\theta}_{M,3}(t) & \dots & c^{\theta}_{M,M-1}(t) & 0 \end{bmatrix}.$$

Here, zeros along the main diagonal mean complete similarity, and the remaining elements are the weights of distinctions between event types, which also generally depend on $t$.

The problem of determining the elements of a distinction weight matrix must be solved based on the substantive aspects of a particular scenario analysis task. For example, in some cases, a logical assumption is that the weight of dissimilarity increases with the scenario deployment time $t$: the closer a given scenario is to the final events of the target scenario, the nearer the assessment of dissimilarity will be to one: $c^{\theta}_{2,1}(t) \to 1$.

**Example 1.** Similarity assessments based on event types.

Let a target scenario $\mathfrak{R}^0$ be defined by the sequence of event types

$TEv^{(\text{sig})0}_m$ = “1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 2, 2, 2, 2, 2, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 6, 6, 6, 6, 6, 6, 6, 7, 7, 7, 7, 7”.

Suppose that two scenarios, $\mathfrak{R}^1$ and $\mathfrak{R}^2$, are obtained by simulations with various control actions:

$TEv^{(\text{sig})1}_m$ = “1, 1, 1, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 8, 8, 8”;

$TEv^{(\text{sig})2}_m$ = “1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 2, 2, 2, 2, 2, 2, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 7, 7, 7, 7, 7”.

According to this example, the scenario $\mathfrak{R}^1$ contains a new event type (type 8), absent in the target scenario. Figure 6 shows the distribution of event types for the scenarios $\mathfrak{R}^0$ and $\mathfrak{R}^1$ by simulation steps; and Fig. 7 presents the corresponding distribution for the scenarios $\mathfrak{R}^0$ and $\mathfrak{R}^2$.

Due to formulas (3) and (4), the measures of similarity are $\delta_H(\mathfrak{R}^0, \mathfrak{R}^1) = 0.417$ and $\delta_H(\mathfrak{R}^0, \mathfrak{R}^2) = 0.806$: the scenarios $\mathfrak{R}^0$ and $\mathfrak{R}^1$ are far from identical, while the scenarios $\mathfrak{R}^0$ and $\mathfrak{R}^2$ coincide by approximately 80%. Consequently, for the two scenarios obtained, the control action $U^2(t)$ that leads to the situation development according to the scenario $\mathfrak{R}^2$ is more effective (closer to the target). Recall that the optimality criterion is the selection of a scenario from the obtained spectrum that is closest to the target one.

Hence, the scenario $\mathfrak{R}^2$ is optimal in the pair of ($\mathfrak{R}^1$, $\mathfrak{R}^2$) and, consequently, strictly preferable ($\mathfrak{R}^2 \succ \mathfrak{R}^1$), i.e.

$$\delta_H(\mathfrak{R}^0, \mathfrak{R}^2) > \delta_H(\mathfrak{R}^0, \mathfrak{R}^1) \Rightarrow \mathfrak{R}^2 \succ \mathfrak{R}^1,\ U^2 \succ U^1.$$

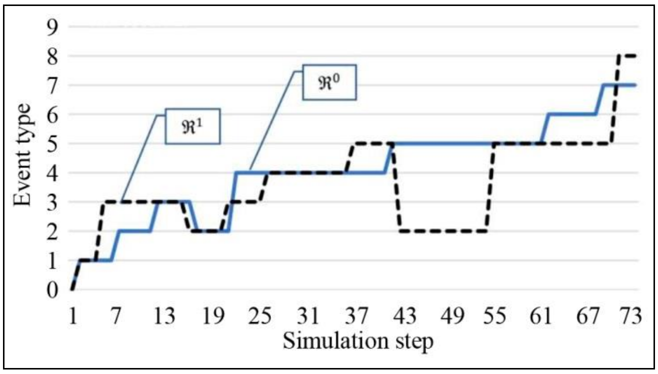



**Fig. 6. An example of overlaying the scenarios $\mathfrak{R}^0$ and $\mathfrak{R}^1$, represented by event types.**

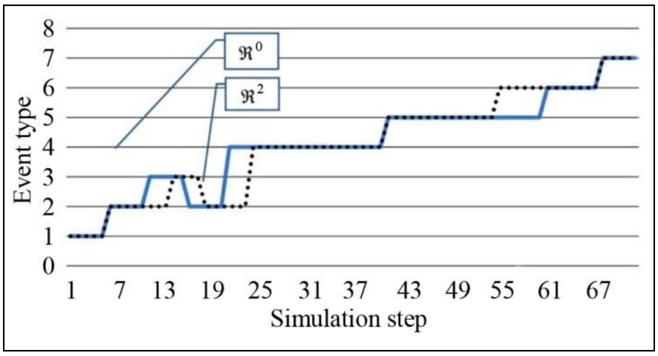



**Fig. 7. An example of overlaying the scenarios $\mathfrak{R}^0$ and $\mathfrak{R}^2$, represented by event types.**

The scenarios $\mathfrak{R}^1$ and $\mathfrak{R}^2$ may also be in a weak preference relation with respect to $\mathfrak{R}^0$ if

$$\delta_H(\mathfrak{R}^0,\mathfrak{R}^1) = \delta_H(\mathfrak{R}^0,\mathfrak{R}^2) \Rightarrow \mathfrak{R}^1 \succsim \mathfrak{R}^2, U^1 \succsim U^2,$$

or

$$\left|\delta_H(\mathfrak{R}^0,\mathfrak{R}^1) - \delta_H(\mathfrak{R}^0,\mathfrak{R}^2)\right| \le \varepsilon \Rightarrow \mathfrak{R}^1 \succsim \mathfrak{R}^2, U^1 \succsim U^2,$$

where $\varepsilon$ is an admissible comparison error.

Suppose that the similarity of scenarios to the target one is determined without fuzziness. Then any two obtained scenarios, $\mathfrak{R}^1$ and $\mathfrak{R}^2$, can be qualitatively represented in terms of their distinction from the target scenario $\mathfrak{R}^0$ using the binary comparison functions (4):

$$\Theta(\mathfrak{R}^0, \mathfrak{R}^1) = \left\{\theta\left(TEv^{(\text{sig})0}(k),\ TEv^{(\text{sig})1}(k)\right)\right\},$$

$$\Theta(\mathfrak{R}^0, \mathfrak{R}^2) = \left\{\theta\left(TEv^{(\text{sig})0}(k),\ TEv^{(\text{sig})1}(k)\right)\right\}.$$

Then

$\Theta(\mathfrak{R}^0, \mathfrak{R}^1)$ = “000111111100001000011111000000 000001111101111111111111000000011111111111”,

$\Theta(\mathfrak{R}^0, \mathfrak{R}^2)$ = “0000000000111001100011100000000 00000000000000000000000001111110000000000000000”, where 0 and 1 indicate no distinction (a distinction, respectively). Since the data under consideration are binary, separators are not required.

By converting zeros to ones and vice versa according to the formula

$$\theta^*\left(TEv^{(\mathrm{sig})0}(k), TEv^{(\mathrm{sig})j}(k)\right) = \begin{cases} 1 \text{ if } TEv^{(\mathrm{sig})0}(k) = TEv^{(\mathrm{sig})j}(k) \\ 0 \text{ if } TEv^{(\mathrm{sig})0}(k) \neq TEv^{(\mathrm{sig})j}(k), \end{cases}$$

we obtain the following representations for assessing the similarity of scenarios to the target scenario:

$\Theta\left(\mathfrak{R}^0, \mathfrak{R}^1\right)^* =$ "111000000001111011110000011111111111100000100000000000001111111000000000000000",

$\Theta\left(\mathfrak{R}^0, \mathfrak{R}^2\right)^* =$ "111111111100011001110001111111111111111111111111111111100000011111111111".♦

## 3. DEFINITION OF THE MEASURE OF CONNECTIVITY OF TWO SCENARIOS

The representations $\Theta\left(\mathfrak{R}^0, \mathfrak{R}^i\right)^*$ and $\Theta\left(\mathfrak{R}^0, \mathfrak{R}^j\right)^*$ can be used to define a measure of connectivity of two scenarios $\mathfrak{R}^i$ and $\mathfrak{R}^j$.

Let us introduce several designations:

- $\theta_i^*(k) = \begin{cases} 1 \text{ if } \theta^*\left(TEv^{(\mathrm{sig})0}(k), TEv^{(\mathrm{sig})i}(k)\right) = 1 \\ 0 \text{ if } \theta^*\left(TEv^{(\mathrm{sig})0}(k), TEv^{(\mathrm{sig})i}(k)\right) = 0; \end{cases}$

- $\theta_j^*(k) = \begin{cases} 1 \text{ if } \theta^*\left(TEv^{(\mathrm{sig})0}(k), TEv^{(\mathrm{sig})j}(k)\right) = 1 \\ 0 \text{ if } \theta^*\left(TEv^{(\mathrm{sig})0}(k), TEv^{(\mathrm{sig})j}(k)\right) = 0; \end{cases}$

- $Nr_I$ and $Nr_J$ are the numbers of ones (matches in the steps of the scenario event types with the target) in the sequences $\Theta\left(\mathfrak{R}^0, \mathfrak{R}^i\right)^*$ and $\Theta\left(\mathfrak{R}^0, \mathfrak{R}^j\right)^*$, respectively, i.e., $Nr_I = \sum_k^K \theta_j^*(k)$ and $Nr_J = \sum_k^K \theta_j^*(k)$;

- $Nr_i$ and $Nr_j$ are the numbers of zeros (mismatches in the steps of the scenario event types with the target) in the sequences $\Theta\left(\mathfrak{R}^0, \mathfrak{R}^i\right)^*$ and $\Theta\left(\mathfrak{R}^0, \mathfrak{R}^j\right)^*$, respectively, i.e., $Nr_i = K - Nr_I$ and $Nr_j = K - Nr_J$, where $K$ is the number of simulation steps;

- $Nr_{IJ}$, $Nr_{iJ}$, $Nr_{Ij}$, and $Nr_{ij}$ are the numbers of pairs (1, 1), (0, 1), (1, 0), and (0, 0) in $\left(\Theta\left(\mathfrak{R}^0, \mathfrak{R}^i\right)^*, \Theta\left(\mathfrak{R}^0, \mathfrak{R}^j\right)^*\right)$, respectively:

$$Nr_{IJ} = \sum_k^K \theta_i^*(k)\theta_j^*(k),$$

$$Nr_{ij} = \sum_k^K \left(1 - \theta_i^*(k)\right)\left(1 - \theta_j^*(k)\right),$$

$$Nr_{Ij} = \sum_k^K \theta_i^*(k)\left(1 - \theta_j^*(k)\right),$$

$$Nr_{iJ} = \sum_k^K \theta_j^*(k)\left(1 - \theta_i^*(k)\right).$$

We calculate the values of these parameters for $\Theta\left(\mathfrak{R}^0, \mathfrak{R}^1\right)^*$ and $\Theta\left(\mathfrak{R}^0, \mathfrak{R}^2\right)^*$ (Table 3).

With the above designations, several well-known methods can be employed to define measures of connectivity [3] in scenario analysis. These methods consider only the qualitative characteristics of scenarios and correspond to the above definition; their values for $\Theta\left(\mathfrak{R}^0, \mathfrak{R}^1\right)^*$ and $\Theta\left(\mathfrak{R}^0, \mathfrak{R}^2\right)^*$ are presented in Table 4.

Thus, for different calculation methods applied to the scenario spectrum, we obtain different assessment matrices $\|c\|$. The rather diverse values obtained in the example (Table 4) can be explained as follows. Some methods take into account the entire initial set of events, whereas others are based on the characteristics of event pairs.

These methods for assessing scenarios can be extended by adding a parameter that characterizes the cost $Cost(U^j)$ of implementing the control actions for the realization of a scenario $\mathfrak{R}^j$. It should also be considered within the range from 0 to 1. In this case, the similarity and costs are considered jointly.

The procedures discussed can be applied in cases with multiple target scenarios $\mathfrak{R}^0 = \left\{\mathfrak{R}^{0,m}\right\}$, $m = 1, \ldots, M$ (e.g., of different priorities), or if preference relations are defined on the set of target scenarios (e.g., strict $\mathfrak{R}^{0,1} \succ \mathfrak{R}^{0,2}$ or weak $\mathfrak{R}^{0,1} \succsim \mathfrak{R}^{0,2}$). In this case, each obtained scenario is compared with each target scenario, and the resulting similarity assessments can be normalized by a priority coefficient or by a preference relation (in practice, in accordance with ordinal utility theory).

If intermediate events are not important, then the goals constitute the final states (events).

*Table 3*

**The intermediate parameters used in assessment formulas**

| Parameter | Symbol in $\Theta\left(\mathfrak{R}^0, \mathfrak{R}^i\right)^*$ | Symbol in $\Theta\left(\mathfrak{R}^0, \mathfrak{R}^j\right)^*$ | Value |
|---|---|---|---|
| $Nr_I$ | 1 | – | 30 |
| $Nr_j$ | – | 1 | 58 |
| $Nr_i$ | 0 | – | 42 |
| $Nr_J$ | – | 0 | 14 |
| $Nr_{IJ}$ | 1 | 1 | 19 |
| $Nr_{iJ}$ | 0 | 1 | 3 |
| $Nr_{Ij}$ | 1 | 0 | 11 |
| $Nr_{ij}$ | 0 | 0 | 39 |

*Table 4*

**Methods for defining the connectivity of scenarios**

| Measure | Formula | Value |
|---|---|---|
| The Jaccard index (similarity coefficient) | $c_{ij} = \frac{Nr_{IJ}}{Nr_{IJ} + Nr_{iJ} + Nr_{Ij}}$ | 0.2753 |
| The simple matching coefficient | $c_{ij} = \frac{Nr_{IJ}}{K}$ | 0.2638 |
| The Tanimoto coefficient | $c_{ij} = \frac{Nr_{IJ}}{Nr_I + Nr_J - Nr_{IJ}}$ | 0.2753 |
| The Sokal–Michener coefficient | $c_{ij} = \frac{Nr_{IJ} + Nr_{ij}}{K}$ | 0.3055 |
| The Rogers–Tanimoto coefficient | $c_{ij} = \frac{Nr_{IJ} + Nr_{ij}}{K + Nr_{iJ} + Nr_{Ij}}$ | 0.1803 |
| The generalized Kulczynski coefficient | $c_{ij} = \frac{1}{4}\left(\frac{Nr_{IJ}}{Nr_I} + \frac{Nr_{IJ}}{Nr_J} + \frac{Nr_{ij}}{Nr_i} + \frac{Nr_{ij}}{Nr_j}\right)$ | 0.3116 |
| The Dice coefficient (the Sørensen–Dice coefficient, F1 score) | $c_{ij} = \frac{2Nr_{IJ}}{2Nr_{IJ} + Nr_{iJ} + Nr_{Ij}}$ | 0.4318 |
| The coincidence weight coefficient | $c_{ij} = \frac{2Nr_{IJ}}{K + Nr_{IJ} + Nr_{ij}}$ | 0.4042 |
| The mismatch weight coefficient | $c_{ij} = \frac{Nr_{IJ}}{Nr_{IJ} + 2(Nr_{Ij} + Nr_{iJ})}$ | 0.1596 |
| The Kulczyinski coefficient | $c_{ij} = \frac{1}{2}\left(\frac{Nr_{IJ}}{Nr_I} + \frac{Nr_{IJ}}{Nr_J}\right)$ | 0.4804 |
| The Ochiai coefficient | $c_{ij} = \frac{Nr_{IJ}}{\sqrt{Nr_I Nr_J}}$ | 0.4554 |
| The generalized Ochiai coefficient | $c_{ij} = \frac{Nr_{IJ} Nr_{ij}}{\sqrt{Nr_I Nr_J Nr_i Nr_j}}$ | 0.0563 |
| The confirmation coefficient | $c_{ij} = \frac{1}{K}\left(Nr_{IJ} + \frac{Nr_{ij}}{K + 1 - Nr_{IJ}}\right)$ | 0.2646 |

A target scenario can also be specified by a nonlinear structure of events, forming a scenario-technological reserve for goal achievement under a threat of certain events. An example is a digraph connecting the initial states of the modeled system ($h_i$) with the final states ($e_i$) (Fig. 8).

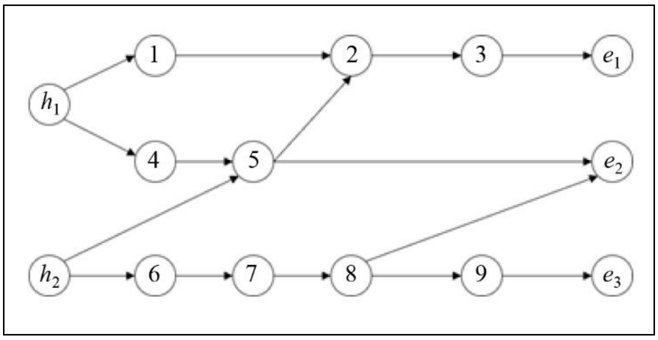



**Fig. 8. An example of a nonlinear target scenario.**

Possible paths between initial and final vertices (events) constitute scenarios for achieving goals.

In the case of a nonlinear target scenario, several analysis tasks arise for the structure corresponding to the graph model, such as finding critical events required to achieve specific or all goals, finding strongly connected components (event cycles), etc. These tasks are not considered here: algorithms for solving them have already been implemented (e.g., in [7]).

Another important application of the above assessment method is to train a graph-based scenario model. This problem can be formulated as tuning weights or refining functional relationships to minimize the error, i.e., the difference between the simulation results and a reference scenario (multiple reference scenarios), e.g., known from historical observations or determined by experts.

## 4. ASSESSING THE SIMILARITY OF SCENARIOS BASED ON EVENT PARAMETERS

In a possible variant of the target scenario, there are goal-significant events at separate stages (time intervals). In this case, the target scenario may be discontinuous (or consist of several noncontiguous fragments): no requirements are imposed on the system's behavior on certain time intervals. For such a "discontinuous" target scenario, it is necessary to introduce an indifferent event type 0. Then the extended alphabet for representing scenarios as words takes the form

$$C^{\mathfrak{R}} = \{\text{“,”}, 0, 1, 2, 3, 4, 5, \ldots\} = \mathbb{N} \cup \text{“,”},$$

$$\theta\left(TEv^{(\text{sig})0}(k),\ TEv^{(\text{sig})j}(k)\right)$$

$$= \begin{cases} 0 \text{ if } TEv^{(\text{sig})0}(k) = TEv^{(\text{sig})j}(k) \\ \qquad \vee TEv^{(\text{sig})0}(k) = 0 \\ 1 \text{ if } TEv^{(\text{sig})0}(k) \neq TEv^{(\text{sig})j}(k). \end{cases}$$

In addition, the similarity of scenarios can be assessed using measures of similarity based on the divergences of the event's *f*-scenarios. For a uniform notation, we will define *f*-scenarios also in the local time (at the *k*th simulation step) of the assessed scenarios. In other words, for a factor $v_n$ in a scenario $\mathfrak{R}^j$, $FSc\left(v_n^j(k)\right) = FSc\left(v_n^j(t - t_r)\right)$.

For example, assume that at the *k*th local step of a scenario $\mathfrak{R}^j$, an expert-significant event is a sequence of *f*-scenarios of significant factors:

$$E^{(\text{sig})j}(k) = \left\{FSc\left(v_n^j(k)\right)\right\},\ v_n^j \in V^{(\text{sig})}.$$

Then we have

$$\theta\left(TEv^{(\text{sig})0}(k), TEv^{(\text{sig})j}(k)\right)$$
$$= \frac{1}{\left|V^{(\text{sig})}\right|} \sum_{v_n \in V^{(\text{sig})}} \theta^{(f)}\left(FSc\left(v_n^0(k)\right), FSc\left(v_n^j(k)\right)\right), \tag{5}$$

where

$$\theta^{(f)}\left(FSc\left(v_n^0(t_i)\right), FSc\left(v_n^j(t_i)\right)\right)$$
$$= \begin{cases} 0 \text{ if } FSc\left(v_n^0(t_i)\right) = FSc\left(v_n^j(t_i)\right) \\ 1 \text{ if } FSc\left(v_n^0(t_i)\right) \neq FSc\left(v_n^j(t_i)\right). \end{cases} \tag{6}$$

Alternatively, let $c\left(d^0,\ d^j,\ t\right) = [0, 1]$ be the weight of the distinction between the types of factor dynamics $d^0$ and $d^j$, where $d$ = 1, 2, 3, 4, 5, 6. In this case,

$$\theta^{(f)}\left(FSc\left(v_n^0(t_i)\right), FSc\left(v_n^j(t_i)\right)\right)$$
$$= c_{FSc\left(v_n^0(t_i)\right),\ FSc\left(v_n^j(t_i)\right)}(t). \tag{7}$$

Thus, a distinction weight matrix is introduced: the main diagonal consists of zeros, while the remaining elements are the weights of the distinctions in factor dynamics (generally, depending on the simulation step $t$):

$$c(t) = \begin{bmatrix} 0 & c_{1,2}(t) & c_{1,3}(t) & c_{1,4}(t) & c_{1,5}(t) & c_{1,6}(t) \\ c_{2,1}(t) & 0 & c_{2,3}(t) & c_{2,4}(t) & c_{2,5}(t) & c_{2,6}(t) \\ c_{3,1}(t) & c_{3,2}(t) & 0 & c_{3,4}(t) & c_{3,5}(t) & c_{3,6}(t) \\ c_{4,1}(t) & c_{4,2}(t) & c_{4,3}(t) & 0 & c_{4,5}(t) & c_{4,6}(t) \\ c_{5,1}(t) & c_{5,2}(t) & c_{5,3}(t) & c_{5,4}(t) & 0 & c_{5,6}(t) \\ c_{6,1}(t) & c_{6,2}(t) & c_{6,3}(t) & c_{6,4}(t) & c_{6,5}(t) & 0 \end{bmatrix}.$$

Theoretically, the distinction weight matrix may be asymmetric with respect to the main diagonal. (Although this would violate the symmetry of the measure of similarity.) Perhaps, in some cases, violating the latter property agrees with the logic of a control problem, but they are not considered here. Therefore, by assumption, the distinction weight matrix is symmetric: $c_{i,j}(t)=c_{j,i}(t)$.

In the general case, there may exist distinction weight matrices for each significant factor $v_n \in V^{(\mathrm{sig})}$, i.e., $c(t, v_n)$. In this case, some elements on the main diagonal are allowed to be nonzero. For example, $c_{2,2}(t, v_3)=1$ means that in a target scenario, a decrease in the factor $v_3$ at a time instant $t$ is undesirable.

Distinctions can also be determined by conditional expressions. For example, if a factor in the compared scenarios demonstrates noncoincident trends of increase and decrease, then for the other factor, the weight between increase and constant value is 0.1, and 0.5 otherwise. That is, the degree of neglecting such a dissimilarity is increased.

Let the weight of the distinction between increase and decrease dynamics be 1, and between increase and constant value be 0.2 for $t=[10, 20]$, etc. In other words, there may be different similarity weights for distinct dynamics on certain time intervals of the compared scenarios, e.g.,

$$\left(FSc\left(v_n^0(t_i)\right)=1, FSc\left(v_n^j(t_i)\right)=3, t\right)$$

$$=c_{1,3}(t)=\begin{cases} 0.2 \text{ for } t=[10, 20] \\ 0.1 \text{ for } t=[31, 45] \\ 1 \text{ otherwise.} \end{cases}$$

Then the measure of similarity $\delta_H$ between scenarios $\mathfrak{R}^0$ and $\mathfrak{R}^j$, which represent finite event sequences $\mathfrak{R}^0=\left\{E_1^{(\mathrm{sig})0}, E_2^{(\mathrm{sig})0},\ldots, E_K^{(\mathrm{sig})0}\right\}$ and $\mathfrak{R}^j=\left\{E_1^{(\mathrm{sig})j}, E_2^{(\mathrm{sig})j},\ldots, E_K^{(\mathrm{sig})j}\right\}$ or finite event types $T\mathfrak{R}^0=\left\{TE_1^{(\mathrm{sig})0}, TE_2^{(\mathrm{sig})0},\ldots, TE_K^{(\mathrm{sig})0}\right\}$ and $T\mathfrak{R}^j=\left\{TE_1^{(\mathrm{sig})j}, TE_2^{(\mathrm{sig})j},\ldots, TE_K^{(\mathrm{sig})j}\right\}$, is defined by the Hamming distance as

$$\delta_H\left(\mathfrak{R}^0, \mathfrak{R}^j\right)=1-\frac{1}{K}\sum_{k=1}^{K}\theta\left(TEv_m^{(\mathrm{sig})0},\ TEv_m^{(\mathrm{sig})j}\right).$$

In the case of similarity assessment based on event components, e.g., as in the relations (5)–(7), $TEv_m^{(\mathrm{sig})j}$ will be a matrix with $\left|V^{(\mathrm{sig})}\right|$ columns (equal to the number of event-significant factors) and $K$ columns (equal to the number of simulation steps):

$$TEv^{(\mathrm{sig})j}=\left\{E^{(\mathrm{sig})j}(k)=\left\{FSc\left(v_n^j(k)\right)\right\}, v_n^j \in V^{(\mathrm{sig})}\right\}.$$

Figure 9 presents the stochastic scenario modeling results obtained by automatically running a scenario model five times; in this model, the weight of the "Canada's influence in the Arctic" arc is specified by a function with pseudo-randomly generated parameters.

In Fig. 9, the top scenario (an event type sequence) is compared with the five scenarios obtained from the simulation, and the percentage of coincidence is calculated.

**Example 2.** Similarity assessments based on *f*-scenarios (by event detail).

Figures 10 and 11 show the matrices of step-by-step changes in the *f*-scenarios of the significant factors corresponding to the two scenarios $\mathfrak{R}^1$ and $\mathfrak{R}^2$ obtained by simulation modeling. (The matrix in Fig. 11 corresponds to the scenario presented in textual event-by-event form in Fig. 4.)

Then, by formulas (3), (5)–(7), for the distinction weight matrix composed $c_{i,j}(t)=1$ for $i \neq j$, and $c_{i,i}(t)=0$, we obtain $\delta_H\left(\mathfrak{R}^1, \mathfrak{R}^2\right)\approx 0.689$.

If the coincidences of events in the last few steps (e.g., the last 20 steps, from 52 to 71) are critically important for assessment, and the distinction weights at the preceding steps grow linearly with the simulation step, i.e.,

$$c_{i,j}(t)=\begin{cases} t/52 \text{ if } i\neq j,\ t=1, 2,\ldots, 52 \\ 1 \text{ if } i\neq j,\ t>52 \\ 0 \text{ if } i=j, \end{cases}$$

then $\delta_H\left(\mathfrak{R}^1, \mathfrak{R}^2\right)\approx 0.748$. ♦

Let $\mathfrak{R}^j \in \breve{\mathfrak{R}}=\left\{\mathfrak{R}^j\right\},\ j=1, 2,\ldots, N^{\mathfrak{R}}$, be a spectrum (set) of scenarios obtained under various conditions, and let $\mathfrak{R}^0$ be a target scenario. Then it is possible to construct a similarity matrix of dimensions $\left[N^{\mathfrak{R}}+1, N^{\mathfrak{R}}+1\right]$ based on the pairwise measures of their similarity:

$$\delta_H = \begin{bmatrix} 1 & \delta_H(1,0) & \ldots & \delta_H(0,\, N^{\mathfrak{R}}-1) & \delta_H(0,\, N^{\mathfrak{R}}) \\ \delta_H(1,0) & 1 & \ldots & \delta_H(1, N^{\mathfrak{R}}-1) & \delta_H(1,\, N^{\mathfrak{R}}) \\ \ldots & \ldots & \ldots & \ldots & \ldots. \\ \delta_H(N^{\mathfrak{R}}-1,0) & \delta_H(N^{\mathfrak{R}}-1,1) & \ldots & 1 & \delta_H(N^{\mathfrak{R}}-1, N^{\mathfrak{R}}) \\ \delta_H(N^{\mathfrak{R}},0) & \delta_H(N^{\mathfrak{R}},1) & \ldots & \delta_H(N^{\mathfrak{R}},\, N^{\mathfrak{R}}-1) & 1 \end{bmatrix},$$

where $\delta_H(i,\ j) = \delta_H\left(\mathfrak{R}^i,\ \mathfrak{R}^j\right)$, $i,\ j = 0, 1, 2, \ldots, N^{\mathfrak{R}}$, is a measure of similarity (or a closeness coefficient) of scenarios $\mathfrak{R}^i$ and $\mathfrak{R}^j$, which is a nonnegative real function such that

$$0 \le \delta_H\left(\mathfrak{R}^i,\ \mathfrak{R}^j\right) \le 1,$$

$$\delta_H\left(\mathfrak{R}^i,\ \mathfrak{R}^i\right) = 1,$$

$$\delta_H\left(\mathfrak{R}^i, \mathfrak{R}^j\right) = \delta_H\left(\mathfrak{R}^j, \mathfrak{R}^i\right).$$


**Fig. 9. Assessment of the obtained scenarios based on their similarity to the target.**

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $v_5$ | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 2 | 5 | 5 | 5 | 5 | 5 | 2 | 2 | 2 | 2 | 2 | 1 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| $v_6$ | 0 | 0 | 0 | 0 | 0 | 3 | 3 | 3 | 3 | 3 | 5 | 5 | 5 | 5 | 5 | 3 | 3 | 3 | 3 | 3 | 1 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| $v_9$ | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 2 | 5 | 5 | 5 | 5 | 5 | 2 | 2 | 2 | 2 | 2 | 1 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| $v_{11}$ | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 2 | 5 | 5 | 5 | 5 | 5 | 2 | 2 | 2 | 2 | 2 | 1 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| $v_{12}$ | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 2 | 5 | 5 | 5 | 5 | 5 | 2 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| $v_{14}$ | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| $v_{15}$ | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 5 | 5 | 5 | 5 | 5 | 1 | 1 | 1 | 1 | 1 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| $v_{19}$ | 0 | 0 | 0 | 0 | 0 | 6 | 6 | 6 | 6 | 6 | 5 | 5 | 5 | 5 | 5 | 6 | 6 | 6 | 6 | 6 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |

| | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 | 61 | 62 | 63 | 64 | 65 | 66 | 67 | 68 | 69 | 70 | 71 | 72 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $v_5$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| $v_6$ | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| $v_9$ | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| $v_{11}$ | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| $v_{12}$ | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| $v_{14}$ | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| $v_{15}$ | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| $v_{19}$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |

**Fig. 10. The matrix of step-by-step changes in the *f*-scenarios of the scenario $\mathfrak{R}^1$.**

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $v_5$ | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 2 | 5 | 5 | 5 | 5 | 5 | 2 | 2 | 2 | 2 | 2 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| $v_6$ | 0 | 0 | 0 | 0 | 0 | 3 | 3 | 3 | 3 | 3 | 5 | 5 | 5 | 5 | 5 | 3 | 3 | 3 | 3 | 3 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| $v_9$ | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 2 | 5 | 5 | 5 | 5 | 5 | 2 | 2 | 2 | 2 | 2 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| $v_{11}$ | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 2 | 5 | 5 | 5 | 5 | 5 | 2 | 2 | 2 | 2 | 2 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| $v_{12}$ | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 2 | 5 | 5 | 5 | 5 | 5 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| $v_{14}$ | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| $v_{15}$ | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 5 | 5 | 5 | 5 | 5 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| $v_{19}$ | 0 | 0 | 0 | 0 | 0 | 6 | 6 | 6 | 6 | 6 | 5 | 5 | 5 | 5 | 5 | 6 | 6 | 6 | 6 | 6 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |

| | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 | 61 | 62 | 63 | 64 | 65 | 66 | 67 | 68 | 69 | 70 | 71 | 72 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $v_5$ | 5 | 5 | 5 | 5 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| $v_6$ | 5 | 5 | 5 | 5 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| $v_9$ | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| $v_{11}$ | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 2 | 2 | 2 | 2 | 2 |
| $v_{12}$ | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| $v_{14}$ | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| $v_{15}$ | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| $v_{19}$ | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |

**Fig. 11. The matrix of step-by-step changes in the *f*-scenarios of the scenario $\mathfrak{R}^2$.**

Note that the above method and scenario assessment algorithm can be further developed based on pairwise comparison. In particular, it is possible to apply methods expanding the assessment capabilities by allowing the use of certain operations over scenario events (Table 5).

For example, such assessments include:

• the Levenshtein distance [9]; for the problem under consideration, the minimum number of event manipulation operations (deletion, insertion, or reversal) required for the event-by-event coincidence of two scenarios. Moreover, if each of these operations has a particular weight, then it is possible to optimize the total weight:

$$\min\left(N^{del}c^{del} + N^{ins}c^{ins} + N^{rev}c^{rev}\right).$$

Here, the Wagner–Fisher algorithm is practical [10];

• the Damerau–Levenshtein distance [7], which is obtained by adding to the list of allowed operations (the definition of the Levenshtein distance) the transposition operation (the permutation of two adjacent events), which may also have a weight:

$$\min\left(N^{del}c^{del} + N^{ins}c^{ins} + N^{rev}c^{rev} + N^{trp}c^{trp}\right);$$

• the longest common subsequence, the Jaro–Winkler similarity [11]; in the assessment problem under consideration, it represents the minimum number of single-event transformations required for matching two assessed scenarios. (Only insertions and deletions, but not replacements, are allowed.)

These and other algorithms [12] are rather simple to implement in software; also, ready-made procedures can be used. For example, Python has a library with various metrics for assessing the similarity of strings; an asymptotic upper bound of their implementation complexity varies from the sum of string lengths to their product.

## CONCLUSIONS

The complexity of solving scenario management problems is that when assessing the effectiveness of scenarios obtained through simulation modeling under certain control actions, one often faces, first, difficulties in identifying events, and second, representing them in a form suitable for measuring their similarity with the events of a target scenario.

In this paper, control effectiveness in scenario modeling has been defined by specifying a target scenario as a discrete set of events and selecting an alternative scenario that best matches a control goal. The control optimality criterion has been defined based on the characteristics of the scenarios corresponding to management options (control options), particularly by selecting, from a set of alternative scenarios, the one closest to a target or safe scenario. In other words, the problem of maximizing a measure of their similarity has been solved. Under resource constraints, it is necessary to use indicators with control costs, which leads to the minimax criterion.

An algorithm has been developed for encoding scenarios as a sequence of unique event types implemented on a sequence of simulation steps.

A method has been developed for the pairwise assessment of a single scenario from the spectrum of scenarios obtained with a target scenario based on the similarity of events at each discrete time instant. Scenarios have been represented in the space of unique types of discrete events. Owing to this approach, modified direct comparison procedures of character strings with the weights of distinctions between event types have been used as a measure of similarity of two scenarios; and this comparison has also been used considering the results of certain operations (insertions, deletions, transpositions) over scenario events with the weights of such operations.

*Table 5*

**Operations over scenario events**

| Operation name | Operation designation | The number of operations | Operation weight |
|---|---|---|---|
| Deletion | $O^{del}$ | $N^{del}$ | $c^{del}$ |
| Insertion | $O^{ins}$ | $N^{ins}$ | $c^{ins}$ |
| Reversal | $O^{rev}$ | $N^{rev}$ | $c^{rev}$ |
| Transposition (the permutation of two adjacent events) | $O^{trp}$ | $N^{trp}$ | $c^{trp}$ |

The proposed methods have been developed into a method for determining the similarity of two scenarios based on event coincidence, e.g., on adjacent or separate time intervals. This method uses the above scenario encoding algorithm and a known algorithm for determining the similarity of text strings based on similar combinations of $m$ adjacent characters ($m$-grams). As a result, it is possible to assess the presence of a significant target sequence of events in the scenario obtained. Thus, the goal itself can be specified as a scenario fragment to be implemented in control actions regardless of its position in the achieved scenario. In addition, the comparison results also generate events that are dynamically considered in the modeling process, e.g., as activity conditions for model substructures, indicators, or events initiating a change in the structure of the controlled object, a change in the control agent, etc.

## REFERENCES


1. *Modeli i metody analiza i sinteza stsenariev razvitiya sotsial'no-ekonomicheskikh sistem* (Models and Methods to Analyze and Design Development Scenarios for Socio-economic Systems), books 1 and 2, Shul'ts, V.L. and Kul'ba, V.V., Eds., Moscow: Nauka, 2012. (In Russian.)
2. Chernov, I.V., Scenario-Cognitive Modeling of Complex Systems Based on Event-Driven Identification of Factor Dynamics, *Control Sciences*, 2023, no. 3, pp. 55–64.
3. Vakhobov, A.A. and Kadyrov, A.L., Determination of the Measure of Closeness of Control Problems, *Bulletin of TSULBP*, 2009, no. 4(40), pp. 38–46. (In Russian.)
4. Shul'ts, V.L., Kul'ba, V.V., Shelkov, A.B., and Chernov, I.V., *Stsenarnyi analiz v upravlenii geopoliticheskim informatsionnym protivoborstvom* (Scenario Analysis in Managing Geopolitical Information Confrontation), Moscow: Nauka, 2015. (In Russian.)
5. Deza, M. and Deza, E., Encyclopedia of Distances, Berlin–Heidelberg: Springer, 2013. URL: https://bayanbox.ir/view/6703722356777834611/Encyclopedia-of-Distances-2013-Michel-Marie-Deza-Elena-Deza-auth..pdf (Accessed September 28, 2024.)
6. Blahut, R.E., *Theory and Practice of Error Control Codes*, Boston: Addison-Wesley, 1983.
7. Cormen, T.H., Leiserson, C.I., Rivest, R.L., and Stein, C., *Introduction to Algorithms*, Cambridge: MIT Press, 2009.
8. Richard, W., Hamming Error-Detecting and Error-Correcting Codes, *Bell System Technical Journal*, 1950, vol. 29, no. 2, pp. 147–160.
9. Levenshtein, V.I., Binary Codes Capable of Correcting Deletions, Insertions, and Reversals, *Soviet Physics Dokl.*, 1965, vol. 10, pp. 707–710.
10. Wagner, R.A. and Fischer, M.J., The String-to-String Correction Problem, *Journal of the ACM*, 1974, vol. 21, no 1, pp. 168–173.
11. Winkler, W.E., String Comparator Metrics and Enhanced Decision Rules in the Fellegi–Sunter Model of Record Linkage, in *Proceedings of the Survey Research Methods Sections*, Alexandria, VA: American Statistical Association, 1990, pp. 354–359.
12. Aggarwal, C.C., *Machine Learning for Text*, Cham: Springer International Publishing, 2022.


*This paper was recommended for publication by A.O. Kalashnikov, a member of the Editorial Board.*

*Received October 29, 2025,*
*and revised January 14, 2026.*
*Accepted January 28, 2026.*

**Author information**

**Chernov, Igor Viktorovich.** Cand. Sci. (Eng.), Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
✉ ichernov@gmail.com
ORCID iD: https://orcid.org/0000-0002-8776-3064

**Cite this paper**
Chernov, I.V., Assessing the Effectiveness of Scenarios Obtained by Simulation Modeling. *Control Sciences* **2**, 51–67 (2026).



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