UDC 004.896

V. B. TROFIMOV[1]*, Associate Professor, Candidate of Engineering Sciences*
I. O. TEMKIN[1]*, Head of Department, Doctor of Engineering Sciences, igortemkin@yandex.ru*
S. V. SOLODOV[1]*, Director of College of Information Technologies and Computer Sciences, Candidate of Engineering Sciences*

[1]*National University of Science and Technology–NUST MISIS, Moscow, Russia*

APPLICATION OF CASE-BASED REASONING IN HAZARD EVALUATION IN COMPLEX PROCESS FLOW CONTROL

The article discusses the Case-Based Reasoning method which enables solving new problems that may arise during decision-making by using or adapting solutions of the similar known problems on the basis of accumulated data and knowledge on past situations or cases contained in a knowledge base. The metrics of similarity between the parameters of a current situation and previous cases, and the methods to retrieve and adapt the cases are described. The case information model used for the process management is presented and exemplified.

The conditions and ranges of efficient case-based reasoning application in the socio-technical system control in case of nonstationary, nonlinear and sluggish processes are discussed.

The authors propose the procedure for searching similar cases using the classical metrics and the Random Forest method, and describe the generalized control of a complex process or an object using the concepts of the industrial internet of things and the case-based reasoning.

Keywords: *control, decision-making, case, metric, library of cases, case-based search of solutions*
DOI: *10.17580/em.2023.02.09*

Introduction

The Case-Based Reasoning Technology (CBR-technology) enables identification of a current situation, finding of a suitable case, using the case for solving the problem or, if necessary, adapting an already known solution to handling the current case.

If the made decisions are not only applicable to a specific situation connected with the condition of an object or a process, and are re-usable, and the main source of knowledge on a problem is the experience, the CBR-technology is efficient [1–4]. The CBR approach involves some assumptions, namely, that:

—similar problems have similar solutions;

—a baseline case can be an example for the similar cases further on;

—it is possible to acquire knowledge through a formal description of cases typical of a control object of any nature;

—there are sufficient optimal (efficient) solutions found by a human-operator for frequent situations;

—there is a large amount of information on external influences. This approach is usually applied in such a problem domain which is impossible to describe adequately and correctly by mathematical modeling.

For limiting the field of search for the target problem solution, it is assumed that the similar cases containing the known solutions much more frequently belong in the same class than in different classes, and that these similar classes form compactly localized sets in the space of cases.

A baseline case contains dozens of characteristics of a control object or of a process which runs in the object. Affinity of a case and a problem situation is evaluated from the pairwise comparison of the characteristics (coordinates). We consider a case as a point in an n-dimensional space having n informative characteristics (coordinates). Selection of the characteristics is an independent problem omitted in this article.

CBR-technology application

The CBR-technology is already used in the applied problem solving in planning, forecasting, classification, optimization and regulation on various scales of control, and also in learning and training of personnel involved in operations/dispatch. Such problems have tendency to reoccur. The periodic processes which are expedient to be automated using the CBR-technology include, for example:

- *converter steelmaking* where the number of melting operations (cases) reaches a few thousands yearly, with making over 400 grades of steel. The information model of this case is described in the works [5–7];
- *coal carbonization* which is a complex object of control, featuring sluggishness, variety of raw materials and produced coke qualities and numerous variables, and requires attendance by high-skilled process engineers who make decisions in difficult conditions;
- *permanent mold casting* [8, 9];
- *machining* [10];
- *thermal treatment* [11].

The CBR method is also used in continuous processes (for instance, in gas conditioning [12], flotation [13], continuous casting [14] and cement production [15–17]).

Regarding gas conditioning, it is proposed in [12] to valuate a situation (normal, emergency, potential emergency), predict it, to give recommendations to maintenance personnel, and to assess the confidence of the type and cause of an accident using a distant CBR system and the inputs as the values of main process variables such as: barometric pressure, gas temperature, rate speeds of rotors of turbines and pumps, temperature difference of gas before and after compressor, vibration, and some others.

The application of the CBR method for predicting accidents in mines, which can cause deaths or environmental disasters, is also a relevant scientific trend [18, 19].

The hybrid processes, which are numerous in mining and metallurgy, combine periodic and continuous operations or stages [20, 21]. The works [22–24] describe the application of the CBR method within the geoinformation systems for solving logical problems connected with spatial relocation, i.e., the problems of transportation planning and quality decision-making in operative management with regard to different technical and economic indexes [25].

For selecting a gold recovery flow process, it is proposed in [26] to include the case model with such characteristics as: the ore type, gold ore grade, gold distribution, gold grain size, content of sulphides, minerals of arsenic sulphide class, minerals of copper sulphide class, minerals of iron sulphide class and content of clay. The similarities between gold ore, concentrates and tailings was assessed in [27]. The methods described in [26, 27] were implemented using open-source software myCBR.

In [28] the CBR system was used to predict fuel consumption per truck per trip and general fuel consumption by a truck fleet per shift or per day in open pit mines. This allows identifying a situation when a truck consumes too much fuel, which makes it possible to decide on corrective measures. The cutback in fuel consumption may result in a great reduction in general expenditures connected with mineral mining. The factors that influence fuel consumption and are entered in an information model of a case are: the type of material hauled by a truck, weather conditions (wind velocity, air temperature and pressure), dead time, truck speed and travel frequency, road grade, payload, one-way distance, cycle duration, wait time, unloading time, basic time, underway time, tones–kilometers per hour, frame torque, torque time and sprung load.

CBR-based control

The case-based reasoning and decision-making cycle (CBR-cycle) consists of certain clearly sequenced stages:

—knowledge base (case library) data mining to find and retrieve a similar case of a set of similar case alternatives for the prevailing problem situation which requires managerial decision-making;

—application of the retrieved case in the current problem solving;

—adaptation (revision) of the case-based solution if it fails to meet the objective;

—preservation of the adapted solution as a new case for the future analogous situations (automated case saving or with attraction of an expert to check conformance with operating procedures—this process can be considered as a system learning process).

The efficiency of a CBR-system [12, 29] and productivity of a search for similar cases depends on a chosen similarity metric. This subjectiveness constitutes a major challenge in using the approach. Let us list some basic metrics.

The Euclidean metric is the frequently used geometrical measure of similarity between a previous case and a current problem situation in the analysis of data represented by a set of informative characteristics in a multi-dimensional space [1]:

$$M_1 = \sqrt{\sum_{i=1}^{n} (X_i^s - X_i^c)^2},$$

where X_i^s is the value of an i-th characteristic of the current situation; X_i^c is the value of an i-th characteristic of the previous case; *n* is the number of characteristics.

The characteristics should possess equal dispersion, independence and weight to assess similarity of situations, phenomena, processes and objects. The metric is applicable to the characteristics determined on the continuous numerical intervals. The quality of the results obtained using this metric greatly depends on the selected normalization algorithm of characteristics. The metric loses efficiency with an increase in the number of the characteristics space dimensions. In this connection, for adding up the weight of the most largely spaced situations and to enhance optimization efficiency owing to simplification of a differentiation procedure, the *squared Euclidean distance* is used.

The Manhattan distance (a partial case of the Minkowski metric of the first order) is closer to the actual distance between a past case and a current situation than the Euclidean space. With this measure used, the influence of runouts reduces as separate great differences caused by the runouts in the Manhattan distance between two points is imaged by different broken curves, and the shortest path is not a single one. The cases which are similarly close to a current situation by the Manhattan metric can be differently distant from the situation by the Euclidean metric. The Manhattan metric is described by the expression below:

$$M_2 = \sum_{i=1}^{n} \left| X_i^s - X_i^c \right| .$$

The Chebyshev distance is the maximum difference between numerical values of characteristics in a previous case and a current situation. This metric may also be useful when it is required to retrieve a case which greatly differs from a current situation by any characteristic. This distance is given by:

$$M_3 = \max_{1 \le i \le n} \left| X_i^s - X_i^c \right| .$$

The exponential metric is used when it is necessary to increase or decrease the weight of a distance between two characteristics. This measure is calculated from the formula below:

$$M_4 = \sqrt[r]{\sum_{i=1}^{n} (X_i^s - X_i^c)^p} ,$$

where *p* and *r* are the parameters set by a designer (if *p* and *r* are equal to two, the metric coincides with Euclidean distance).

The Zhuravlev metric is applied to the quantity and quality (e.g., nominal and ordinal) informative characteristics. Efficiency of this measure follows from the possibility of the contensive justification of a threshold ε set by an expert, and allows transiting between the quantity and quality scales. Evidently, some useful information can be lost in this case. The Zhuravlev metric is described by the following system:

$$M_5 = \frac{1}{n}\sum_{i=1}^{n} q_i \ ;$$

$$q_i = \begin{cases} 1, \text{если } \left| X_i^s - X_i^c \right| < \varepsilon_i; \\ 0, \text{если } \left| X_i^s - X_i^c \right| \ge \varepsilon_i, \end{cases}$$

where ε_i is the threshold of an i-th characteristic (for quality characteristics, $q_i = 1$ if their gradations coincide, otherwise, $q_i = 0$).

The Canberra distance is a weighted version of the Manhattan distance, and is determined using one of the methods below:

$$M_6^{I} = \frac{\sum_{i=1}^{n}\left| X_i^s - X_i^c \right|}{\sum_{i=1}^{n} X_i^s + X_i^c}, \quad M_6^{II} = \frac{\sum_{i=1}^{n}\left| X_i^s - X_i^c \right|}{\sum_{i=1}^{n}\left| X_i^s \right| + \left| X_i^c \right|}, \quad M_6^{III} = \frac{\sum_{i=1}^{n}\left| X_i^s - X_i^c \right|}{\sum_{i=1}^{n}\left| X_i^s + X_i^c \right|} .$$

The Bray–Curtis metric may be given by:

$$M_7 = \frac{\sum_{i=1}^{n} 2\min(X_i^s, X_i^c)}{\sum_{i=1}^{n} X_i^s + \sum_{i=1}^{n} X_i^c}$$

[30] (before using this measure, minimax normalizing of values of characteristics is usually carried out).

The listed metrics are only a part of the family of the exiting criteria to evaluate closeness of vectors in the space of characteristics, but are used most frequently.

The choice of an algorithm to retrieve cases from a knowledge base depends on the structure of the latter and on the form of representation and method of storage of cases. These methods may be built on a relational database. The most often forms of representation (structuring) of cases are: the parametric representation (a set of measured values of the case parameters, i.e. a record in the database), or the object-oriented representation in the form of a graph (semantic network), frame, tree, or a predicate.

The repeated use of experience (field data) makes it possible to cut down the time spent to solve a problem, and to enhance the control efficiency. A library of cases is a structured representation of the accumulated experience in the form of an aggregate of data and knowledge. It can be a part of a knowledge base but is usually a CBR-subsystem, which affects the time of finding and retrieving cases.

The aggregate of data and knowledge on methods to described cases and on procedures to manipulate them shape a certain information model. A case information model used in control includes such components as: description of a problem situation (values of characteristics, status of a control object), description of a problem solution (sequence of implemented control actions, control programs), description of a result of the solution and its efficiency (values of quality indexes of the control object functioning). The model may contain references to other cases.

With a view to abating the impacts of stochastic and nonstationary rough noise in the complex socio-technical systems, it is possible to use the algorithms of moving average or exponential first-order or second-order shooting, robust smoothing algorithms, or algorithms of median evaluation, on-off exponential smoothing and median-exponential smoothing. Rough noise may appear because of measuring system fails, nonrepresentative samples or errors in manual data input.

The use of a complex-structure representation of cases requires using a metric and computation procedure of the same complexity. Selection of an adequate metric agreeable with the compactness hypothesis is the most labor-intensive and the least investigated task; it depends on the characteristics of data being processed, objectives of a user, and on the knowledge of a designer on a problem domain. If the similarity between a complex problem situation and the cases, evaluated using the chosen metric, is much less than a threshold set by an expert, the case is created using standard methods and procedures which require much time and resources.

The nearest neighbor analysis may appear ineffective if the measured values of parameters (characteristics) contain fluctuation noises and rugged run-outs, are incomplete, or it is problematic to select a solution when there are a few cases (from the same or some different classes) which are equally distant from the current problem situation [29, 31]. This method requires a large-volume memory as the whole library of cases is used for searching for a solution.

It is advisable to make a decision using a few closest-spaced points (cases) rather than one point, as this enables reduction in the influence of random run-outs, noise and errors which are always present in the field data (the maximum number of the closest-spaced cases is usually not more than 20; the violation of this term often suppresses efficiency of the system because of inclusion of many less similar cases).

With large knowledge bases, for shortening search time, it is recommended to use the case retrieval method based on the decision-making tree where the lead nodes conform with one or a few cases [31]. The choice of a limb of the tree for searching solutions is carried out using the data on a current problem situation. When a leaf node contains a few cases, the nearest neighbor analysis is then used [32]. In larger libraries of cases, efficiency of case retrieval usually lowers. If a case exists for a long time but remains unused and is unadaptable, it should be placed in an archive.

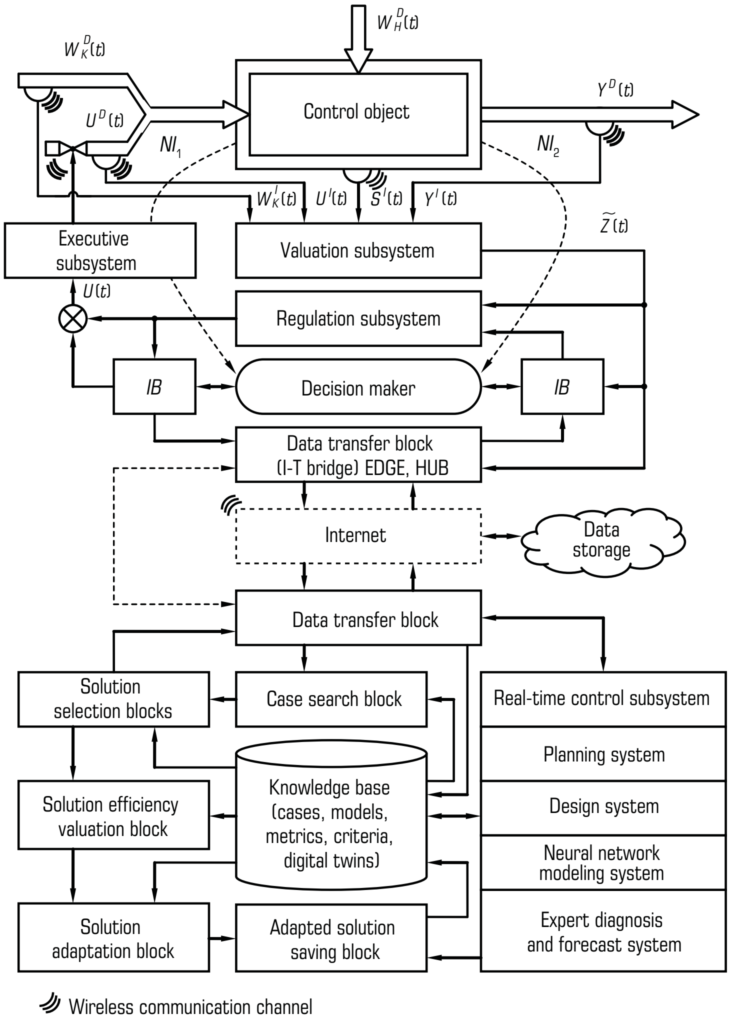

Fig. 1. Structure of CBR-based control

The method of fast retrieval of cases on the basis of their semantic indexation in real time, assuming the use of weights of parameters (characteristics) and additional knowledge on the application domain, is employed in case of small knowledge bases. The coefficients of the relative values of characteristics (for instance, from 0 to 1) are set by an expert or by a group of experts on the basis of their ranking or pair-wise comparison [33].

According to the concept of industrial internet of things, a bulk of control functions may be given to high-performance servers at data processing centers, while the number and complexity of lower level facilities may jump [12, 32]. In the framework of this concept and CBR approach, we propose the structure of a control system in **Fig. 1**. The legend in the figure is as follows: $W_K^D(t)$ —actual controllable external effects at the time t; $W_H^D(t)$ —uncontrollable external effects; $U^D(t)$ —control actions (continuous, period, nonrecurrent); $Y^D(t)$ and $S^D(t)$ —output effects and statuses of an object; $W_K^И(t)$, $U^И(t)$, $Y^И(t)$ and $S^И(t)$ —measured values of variables of a control object; $\widetilde{Z}(t)$ —valuations of variables of the control object; IB—interface block (natural language interface); NI—non-instrumental information.

It is possible to enhance efficiency of control by using digital twins of an object. If the control of a complex process includes collection of data by a set of distributed information-and-measurement systems, there exists natural decomposition of raw data sources, the space contains variable types of characteristics (Booleans, integrals, continuous, images), or a problem situation features a high dimensionality of a space of characteristics, it is proposed to use a few different metrics jointly (**Fig. 2**).

The primary sets of the nearest cases $P_{M1}(Sit_j)$, $P_{M2}(Sit_j)$, …, $P_{MN}(Sit_j)$ for solving a j-th problem situation Sit_j are formed using the corresponding metrics $M_1, M_2, \dots, M_N$. The final set of the cases $P(Sit_j)$ is generated according to the following rule:

$$P(Sit_j) = \{P_{M1}(Sit_j) \cap P_{M2}(Sit_j)\} \cup \{P_{M2}(Sit_j) \cap P_{M3}(Sit_j)\} \cup$$
$$\cup \{P_{M1}(Sit_j) \cap P_{M3}(Sit_j)\} \cup ,\dots, \cup \{P_{MN-1}(Sit_j) \cap P_{MN}(Sit_j)\}.$$

The final set only contains the cases which have the same class with the result of situation recognition. The most frequent cases have the higher values.

The weight of a j-th metric ($j = 1, 2, \dots, N$), used to solve K problems, is given by:

$$w_j = \frac{\sum_{k=1}^{K} r_k}{K},$$

where $r_k = \begin{cases} 1 \text{ } if \text{ solution of a } k\text{-th problem is found using a } j\text{-th metric;} \\ 0 \text{ } if \text{ proposed solution of a } k\text{-th problem needs adaptation.} \end{cases}$

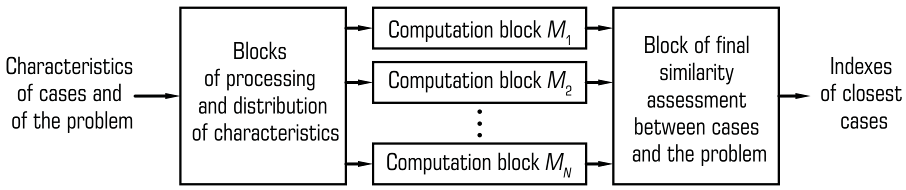

Fig. 2. Structure of search block for close cases

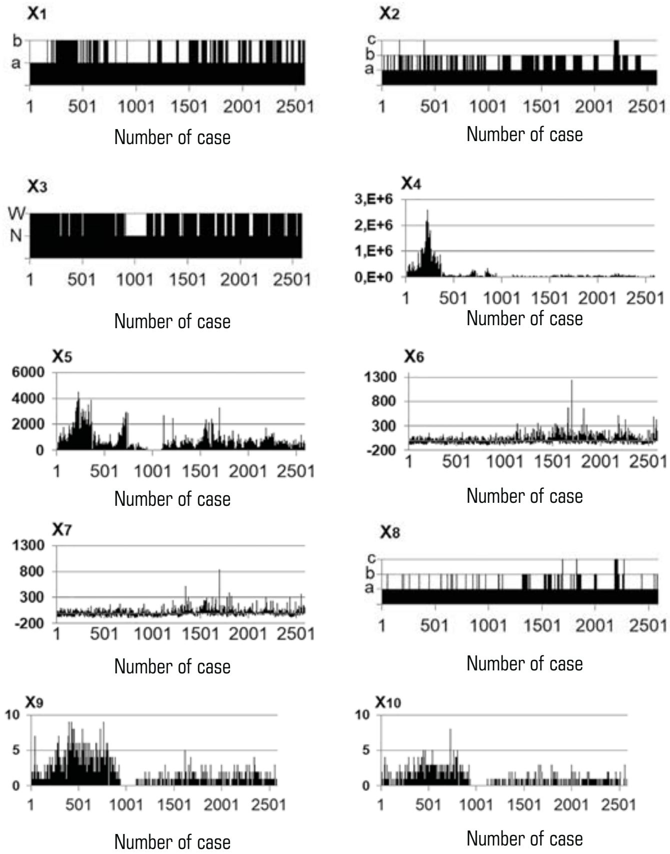

Fig. 3. Diagrams of characteristics from knowledge base
Номер прецедента – Number of case

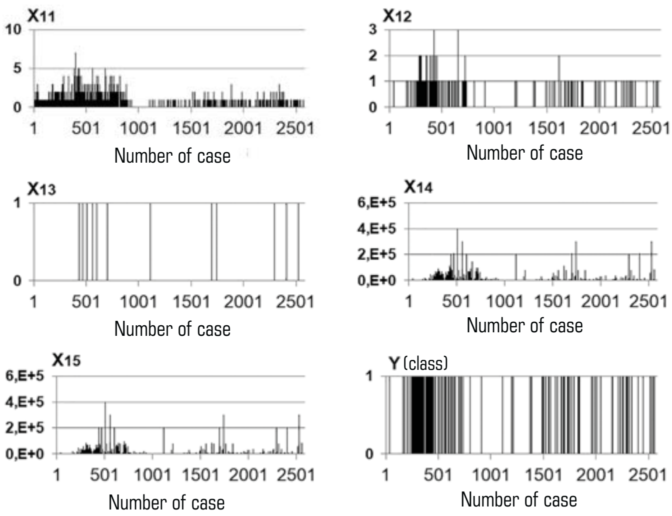

Fig. 4. Diagrams of characteristics and class number

The solution adaptation block uses the methods described below.

Adaptation by substitution consists in *substitution* of some elements (parameters) in the solution of a found case which should be very close to a problem situation. To this end, we find dependences between the parameters of the current situation and the parameters of the appropriate solution in the library of cases.

Adaptation of transformation enables *removing* and *adding* elements in a solution, i.e. implementing its *reorganization*, resequencing of operations by certain rules and using proper operators based on the analysis of differences between the new and retrieved cases. When using adaptation operators which implement searching with return at a branch point, it is necessary to take into account the time of their work with the specific field data with the assumption that the number of possible variants is not high. The repeated use of elements of a solution is excluded.

Case-based reasoning in seismic hazard prediction in coal mines

Coal mines are the complex engineering systems which involve technological impacts on different geological, physicochemical, aerological and other processes that, if uncontrolled, can lead to casualties and accidents.

For solving the problem connected with forecasting seismic hazard (seismicity-induced bump with an energy higher than 10^4 J within the coming eight-hour shift, which means that the prediction interval is eight hours), we used the proposed approach and the field data (**Figs. 3** and **4**) obtained from the monitoring systems in the test coal mine and aggregated within the structured data array Seismic Bumps (from digital library UCL).

The legend in Figs. 3 and 4 is as follows:

X_1—resultant hazard from seismic method;

X_2—resultant hazard from acoustic method;

a—no hazard; b—low hazard; c—high hazard;

X_3—information about shift (W—coal mining, N—development work);

X_4—seismic energy recorded during previous shift by the most active geophone (GMax) out of seismic receivers installed in longwall;

X_5—number of impulses recorded during previous shift by GMax;

X_6—deviation of energy recorded during previous shift by GMax from average value of energy recorded during eight previous shifts;

X_7—deviation of number of impulses recorded during previous shift by GMax from average number of pulses recorded during eight previous shifts;

X_8—resultant shift-wise hazard evaluated by seismic and acoustic methods from recordings by GMax only;

X_9—number of seismic bumps recorded during previous shift;

X_{10}—number of seismic bumps in energy range [10^2–10^3 J] from previous shift;

X_{11}—number of seismic bumps in energy range [10^3–10^4 J] from previous shift;

X_{12}—number of seismic bumps in energy range [10^4–10^5 J] from previous shift;

X_{13}—number of seismic bumps in energy range [10^5–10^6 J] recorded in the last shift;

X_{14}—total energy of seismic bumps recorded during previous shift;

X_{15}— energy of seismic bumps recorded during previous shift;

Y (class)—attribute of solution: 1 means that a high energy seismic bump takes place in the next shift (hazard), 0 means that no high energy seismic bumps take place in the next shift (no hazard) [34].

Factors that may induce seismic bumps are variable, can be expressed quantitatively or qualitatively, i.e. are described using different-nature scales, have measurement errors, their distribution differs from normal and their inter-connections are nonlinear. Some factors result from aggregation of partial indexes, which usually results in a partial loss of useful information. The seismic and acoustic methods fail to provide a high-accurate prediction of seismic hazards in coal mines [34].

The numbers of the closest cases ordered using different metrics, and the related classes (0 or 1) for a test situation $Sit_1 = \{X_1=a;\ X_2=a;\ X_3=W;\ X_4=207930;\ X_5=614;\ X_6=-6;\ X_7=18;\ X_8=a;\ X_9=2;\ X_{10}=2;\ X_{11}=0;\ X_{12}=0;\ X_{13}=0;\ X_{14}=1000;\ X_{15}=700;\ Y=0\}$ are presented as follows:

by the Euclidean metric and square of the Euclidean distance—{(58;0), (102;0), (843;0), (2075;0), (1687;1), (846;0), (493;0), (2022;0)};

by the Manhattan metric—{(58;0), (2075;0), (843;0), (102;0), (1687;1), (493;0), (2022;0), (846;0)};

by the Chebyshev metric—{(58;0), (102;0), (846;0), (843;0), (570;1), (573;0), (493;0), (2002;1)};

by the exponential metric (at $p = 4$, $r = 6$)—{(58;0), (102;0), (843;0), (846;0), (493;0), (573;0), (2075;0), (570;1)};

by the Zhuravlev metric—{(273;0), (274;1), (579;0), (661;0), (681;0), (781;0), (875;0), (2075;0)};

by the Canberra and Bay–Curtis metrics—{(58;0), (2075;0), (102;0), (843;0), (1687;1), (493;0), (846;0), (2022;0)}.

The numbers of the closest cases and the related classes for a test situation $Sit_2 = \{X_1= a;\ X_2= a;\ X_3= N;\ X_4=384230;\ X_5=751;\ X_6=4;\ X_7=6;\ X_8= a;\ X_9=3;\ X_{10}=1;\ X_{11}=2;\ X_{12}=0;\ X_{13}=0;\ X_{14}=9700;\ X_{15}=6000;\ Y=1\}$ are presented as follows:

by the Euclidean metric and square of the Euclidean distance—{(141;0), (388;0), (347;0), (398;1), (699;0), (640;1), (465;1), (445;1)};

by the Manhattan metric—{(445;1), (640;1), (141;0), (398;1), (436;0), (388;0), (646;0), (507;0)};

by the Chebyshev metric—{(141;0), (347;0), (699;0), (865;0), (388;0), (398;1), (465;1), (490;0)};

by the exponential metric (at $p = 4$, $r = 6$)—{(141;0), (347;0), (699;0), (388;0), (398;1), (465;1), (865;0), (490;0)};

by the Zhuravlev metric—{(747;0), (865;0), (2133;0), (141;0), (216;0), (388;0), (398;1), (403;0)};

by the Canberra and Bay–Curtis metrics—{(445;1), (640;1), (141;0), (398;1), (436;0), (388;0), (646;0), (507;0)}.

Before using the metrics, we undertook minimax normalizing of all quantity characteristics (factors). The Zhuravlev metric was applied to the quantity and quality characteristics, and the other metrics were only applied to the quantity characteristics.

If the found primary sets of cases contain more than two classes 1, a hazardous situation is possible. The knowledge base only contains 2584 cases out of which 170 case have class 1. The out-of-balance distribution of cases per classes greatly complicates the procedure and reliability of reasoning as the elements in the ranked sequences of cases have noticeably different weights, Nonetheless, it is evident that if the closest (the first in the sequences) elements by any test metric contain class 1 of a hazardous situation, it is better to initiate the confirmation procedure. To this effect, the authors propose to use the Random Forest Classifier. The Random Forest method is a classifier that contains an ensemble of decision trees. For 100 trees which have a max_depth parameter equal to 2, the classification accuracy was 94.54% from the method of 10-fold cross check (procedure ms.cross_val_score). For 21 trees with the depth parameter of 3 or 4, the classification accuracy grew to 100%. The same accuracy is ensured by a forest of 6 trees with the depth parameters from 5 to 9, or by a forest of 8 trees with the depth parameter from 10 to 15. The hazard prediction by this method uses all test characteristics (see figs. 3 and 4).

The final set of the closest cases required for an object control includes the cases from the primary sets, if their class number agrees with the result of recognition by the Random Forest. In case that a hazardous situation is predicted, the rockburst hazard reduction requires undertaking certain technical and technological procedures.

The development of the procedure for searching close cases used the library of the high-level language Python, which enabled implementation of some basic methods of machine learning—Scikit-learn with regard to the architectural concepts consummated on the spatial platforms CAT-CBR and myCBR.

Conclusions

Automation of argumentation by means of the case-based reasoning aims at creation of libraries of cases for their storage and automated search of their attributes and methods to evaluate similarity between a previous case and a new target situation for selecting the most suitable case, its adaptation techniques and visualization tools.

The known metrics of similarity (closeness) are analyzed. The structure of control using the industrial internet of things and the case-based reasoning is proposed. The advantage of the CBR-technology capable to process both numerical and symbolic information consists in its capacity to solve ill-structured and ill-formalized problems subject to insufficient knowledge on a control object and on its external environment.

The CBR-technology allows direct application of accumulated experience of problem solving, heuristic evaluation, reduction of time of complex problem solving, elaboration of control actions to reach the objective under uncertainty, warning of a user on possible failure, elimination of potential re-use of an erroneous decision and creation of corporate memory and information resource of an enterprise. The scope of saving embraces cases which have both positive and dubious or negative outcome.

The proposed approach differs from other techniques of using knowledge bases by employment of a set of metrics to search for the closest cases based on various quantitative and qualitative information characteristics in real time, application of the Random Forest method, and by incorporation of knowledge bases with information models of cases, descriptions of situations, sets of control actions and their implementation results.

Nevertheless, the topical objectives of the further basic and applied research remain:

—the provision of high productivity of the CBR system at high number of cases;

—the improvement of the validity of expert knowledge on a problem domain for the case description;

—the creation of applications ensuring efficient use of the CBR-technologies for planning behavior of dynamic media.

References

1. Osipov G. S. Methods of Artificial Intelligence. Moscow : Fizmatlit, 2015. 295 p.
2. Sokolov I. A. Theory and practice of application of artificial intelligence methods. *Vestnik RAN*. 2019. Vol. 89, No. 2. pp. 115–119.
3. Vassilyev S. N., Novikov D. A., Bakhtadze N. N. Intelligent control of industrial processes. *IFAC Proceedings Volumes (IFAC-Papers Online)*. 2013. Vol. 46, No. 9. pp. 49–57.
4. Okhtilev M. Yu., Sokolov B. V., Yusupov R. M. Intelligent technologies for monitoring and controlling the structural dynamics of complex objects. Moscow : Nauka, 2006. 410 p.
5. Gu M.-Q., Xu A.-J., Yuan F., He X.-M., Cui Z.-F. An improved CBR model using time-series data for predicting the end-point of a converter. *ISIJ International*. 2021. Vol. 61, No. 10. pp. 2564—2570.
6. Trofimov V. B. An approach to intelligent control of complex industrial processes: An example of ferrous metal industry. *Automation and Remote Control*. 2020. Vol. 81, No. 10. pp. 1856–1864.
7. Cheng Y., Xing J., Dong J., Wang Z., Wang X. Improved CBR for endpoint carbon content prediction of BOF steelmaking. *The 9th International Conference on Intelligent Control and Information Processing (ICICIP)*. 2018. pp. 168–175. DOI: 10.1109/ICICIP.2018.8606688

8. Khosravani M. R., Nasiri S. Injection molding manufacturing process: review of case-based reasoning applications. *Journal of Intelligent Manufacturing*. 2020. Vol. 31. pp. 847–864.
9. Bolender T., Burvenich G., Dalibor M., Rumpe B., Wortmann A. Self-adaptive manufacturing with digital twins. *Proceedings—The 2021 International Symposium on Software Engineering for Adaptive and Self-Managing Systems (SEAMS)*. 2021. pp. 156–166. DOI: 10.1109/SEAMS51251.2021.00029
10. Xu L., Huang C., Li C., Wang J., Liu H. et al. A novel intelligent reasoning system to estimate energy consumption and optimize cutting parameters toward sustainable machining. *Journal of Cleaner Production*. 2020. Vol. 261. DOI: 10.1016/j.jclepro.2020.121160
11. Hadjiski M., Deliiski N., Tumbarkova N. Intelligent hybrid control of thermal treatment processes of wood. *The 2020 IEEE 10th International Conference on Intelligent Systems, IS 2020 – Proceedings*. 2020. pp. 482–489.
12. Kuzyakov O. N., Andreeva M. A. Applying case-based reasoning method for decision making in IIoT system. *The 2020 International Multi-Conference on Industrial Engineering and Modern Technologies, FarEastCon*. 2020. DOI: 10.1109/FarEastCon50210.2020.9271301
13. Yan H., Wang F., Yan G., He D. Hybrid approach integrating case-based reasoning and Bayesian network for operational adjustment in industrial flotation process. *Journal of Process Control*. 2021. Vol. 103. pp. 34–47.
14. Su W., Lei Z. Rough case-based reasoning system for continues casting. *Proceedings of SPIE—The International Society for Optical Engineering*. 2018. Vol. 10696. DOI: 10.1109/IS48319.2020.9200096
15. Li Q., Lu S., Yu H., Wane X., Liu H. Research on optimal setting of cement decomposing furnace temperature based on case reasoning. *Proceedings—The 2020 35th Youth Academic Annual Conference of Chinese Association of Automation, YAC*. 2020. pp. 626–631.
16. Zaytseva E. V., Medyanik N. L. Automated integrated production and selling planning at processing plant in the cement industry. *GIAB*. 2022. No. 2. pp. 111–123.
17. Zaytseva E. V. Strategic management in the cement industry. *GIAB*. 2019. No. 2. pp. 214–220.
18. Li Y., Shao L., Liu S. Case-based reasoning forecast method of coal mine emergencies. Liaoning Gongcheng Jishu Daxue Xuebao (Ziran Kexue Ban). *Journal of Liaoning Technical University (Natural Science Edition*). 2014. Vol. 33, No. 7. pp. 902–906.
19. Kupriyanov V. V., Temkin I. O., Bondarenko I. S. Study of the time characteristics for emergency situations in the coal mines. *Occupational Safety in Industry*. 2022. Vol, No. 1, pp. 39–45.
20. Jinsheng S., Haigang F. Implementation of CBR strategy for combustion control of blast furnace stoves. *Chinese Control and Decision Conference*. 2008. DOI: 10.1109/CCDC.2008.4597517
21. Qiu H., Tian J., Jian L. Parameter fusion modeling method for hot strip rolling process and its application. *Proceedings of the 33rd Chinese Control and Decision Conference*. 2021. Vol. 10696. pp. 374–378. 22. Belyakov S. L., Bozhenyuk A. V., Belyakova M. L. Case based analysis of geo-information models of logistics projects. *Nauka i tekhnologii zheleznykh dorog*. 2019. Vol. 3, No. 2(10). pp. 53–63.
23. Belyakov S. L., Bozhenyuk A. V., Belyakova M. L., Zubkov S. A. Case based reasoning in intelligent geographic information systems for the management of logistics projects. *CEUR Workshop Proceedings: 2nd International Scientific and Practical Conference on Fuzzy Technologies in the Industry—FTI*. 2018. pp. 1–10.
24. Gordienko L. V., Samoylov L. K. Geoinformation model for definition of proximity of logistic precedents. *Sovremennye innovatsionnye tekhnologii podgotovki inzhenernykh kadrov dlya gornoi promyshlennosti i transporta*. 2016. No. 1(3). pp. 294–299.
25. Bondarenko I. S. Elaboration of plans–forecasts based on engineering-and-economic performance of mines. *GIAB*. 2022. No. 3 pp. 97–107.
26. Rintala L., Leikola M., Sauer Ch., Aromaa J., Roth-Berghofer T., Forsén O., Lundström M. Designing gold extraction processes: Performance study of a case-based reasoning system. *Minerals Engineering*. 2017. Vol. 109. pp. 42–53.
27. Leikola M., Sauer C., Rintala L., Aromaa J., Lundström M. Assessing similarities between gold ores, concentrates, and tailings with case-based reasoning. *Minerals Engineering*. 2020. Vol. 146. ID. 106113.
28. Tadubana G., Sigweni B., Suglo R. Case-based reasoning system for prediction of fuel consumption by haulage trucks in open-pit mines. *International Journal of Electrical and Computer Engineering*. 2021. Vol. 11, No. 4. pp. 3129–3136.
29. Biswas S., Devi D., Chakraborty M. A hybrid case based reasoning model for classification in internet of things (IoT) environment. *Journal of Organizational and End User Computing*. 2018. Vol. 30, No. 4. pp. 104–122.
30. Rodionov I. V., Sozontov A. N. On confidence estimation based on quantitative similarity coefficients. *Automation and Remote Control*. 2020. Vol. 81, No. 2. pp. 320–332.
31. Temkin I. O., Klebanov D. A., Deryabin S. A., Konov I. S. Building an intelligent geoinformation system of a mining enterprise using predictive analytics methods. *GIAB*. 2020. No. 3. pp. 114–125.
32. Vagin V. N., Golovina E. Yu., Zagoryanskaya A. A., Fomina M. V. Reliable and plausible inference in intelligent systems. Moscow : Fizmatlit, 2008. 712 p.
33. Luger G. F. Artificial Intelligence: Structures and Strategies for Complex Problem Solving. USA : Pearson Education, 2008. 754 p.
34. Kabiesz J., Sikora B., Sikora M., Wrobel L. Application of rule-based models for seismic hazard prediction in coal mines. *Acta Montanistica Slovaca*. 2013. Vol. 18, No. 4. pp. 262–277.