UDC 004.56+004.67

COMPUTATION OF PROVIDING SERVICES INTEGRAL QUALITY INDEX

Globa L.S., Svetsynska I.S., Volvach I.A.

The National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"

Kyiv, Ukraine

Background. The paper is devoted to review and improvement of existing algorithms for assessing the service provision quality by a telecom operator. Modern approaches to determining the level of service quality in the operator's telecom system require the use of complex mathematical methods and approaches that have significant computational complexity. The proposed approach to assessment of the provided services quality is based on a generalized quality function defined as the geometric average of the individual quality indicators of the operator's telecom monitoring system and allows the formation of fuzzy knowledge base in the form of the structured rules. To obtain the value of the integral indicator the rules of the fuzzy knowledge base are trained using the desirability function and the clustering method. The data set presentation to obtain from the operator's telecom monitoring system at certain time intervals in the form of the fuzzy knowledge base structured rules allows to reduce the making decision time on the service provision quality significantly, as well as to reduce the computational complexity of the service provision quality determining.

Objective. Improving the service provision quality to the end user through "soft" condition control of the operator's system performance indicators and reducing computational complexity in determining their quality.

Methods. The study was carried out based on a large number of literary sources analysis, the theory of fuzzy logic, clustering methods with using the generalized quality function, the theory of fuzzy knowledge base.

Results. An approach is proposed to determining the integral quality indicator of the provided by the telecom operator to the end user services, obtaining the complex non-structured data estimation based on one integral fuzzy indicator and forming on its basis a knowledge system represented as the fuzzy knowledge base. When forming a fuzzy knowledge base, it takes time to learn its rules, but this is compensated by computational load reduction on the system during its operation.

Conclusions. The presented research indicates the need to improve the modern telecom operator's platform for the services provision that realized by an additional block of "soft" services quality management with the help of fuzzy knowledge base. This modification allows instead of processing Big Data from the telecom operator's monitoring system to determine the quality of services based on the rules of the fuzzy knowledge base.

Keywords: Big Data; fuzzy knowledge base; fuzzy logic; service quality; integral quality index.

Introduction

At the present development stage of information and communication technologies, the term Big Data means a number of approaches, tools and methods of processing of structured and unstructured vast amounts of data and their considerable diversity.

Big Data is a term used to identify data sets. It can't be coped with such data sets using existing methodologies and software because of their large size, rate of arrival, analysis and complexity. The researchers as M. Hilbernt, S. Strinivasa and others developed methods and software tools for data transmission and information granules mining from Big Data (objects collections formation, that are usually formed for numeric attributes. They are placed side by side because of their similarity, functional or physical commonality), but the appearance of new data formats requires constant expansion and improvement of existing methods and data analysis tools [1].

With that accumulating data from network nodes rapidly increase every year. This causes necessity in powerful constantly increasing computing resources to increase processing speed and data access. In this regard, necessity to improve large amounts of data processing algorithms becomes more relevance.

According to research were conducted by a number of leading companies in the world [2] telecom operators are faced with the urgent need the complex account of different characteristics of provided services (technical, economic, experience of using services by end-user) due to the rapidly growing range of services and the transition to digital space. They want a clear understanding and process management that occurring between the operator and its subscribers. This whole range of parameters is too large and complex data for the

collecting, processing and analysis with the use of current computing infrastructure and is characterized by:

- significant amount of data (from terabytes to petabytes);
- necessity for high processing speed in real time to reduce the volume of storage;
- heterogeneity (can be structured, unstructured, semi-structured);
- necessity to fulfil a validity requirement (may be disrupted due to the variety of data sources and processing methods, violations of safety requirements);
- importance (using of forecasted methods and analysis methods allows to predict the direction of companies development).

Telecommunication companies are investing a lot of money in analytical tools and services development. With this data telecom operators aim to:

- increasing sales;
- assuring revenue (detecting and preventing revenue leakage);
- reducing churn and fraud;
- improving risk management;
- decreasing operational cost;
- improving visibility into core operations, internal processes and market conditions;
- discerning trends and establishing forecasts;
- cross-selling/up-selling products and service plans.

At the same time for such kind of tasks data analysis is often performed based on data that obtained as a result of the economic operator's activity or sociological interrogations or technical parameters of the operator's infrastructure functioning (for example, the decision to invest in one or another part of the system does not consider the influence and analysis of all possible factors and consequences). For example, if we consider the problem of users' satisfaction degree by services provided by telecom operator, it is quite obvious that the frequent technical failures affect satisfaction degree, price operator's policy and the service performance jointly impact on the final services performance evaluation by subscriber.

Modern facilities of Big Data analysis require a transition from unstructured to structured data, thus forming, relatively speaking, "volume data compression volume to their meaning" and generating data processing strategy for Big Data as "data - information - knowledge - prediction" (Fig. 1.). In this case the entered processing steps are understood as [2]:

- data – streams of raw facts such as business transactions;
- information – clusters of facts that are meaningful to human beings such as making decisions;
- knowledge – data/information organized to convey understanding, experience, accumulated learning, and expertise;
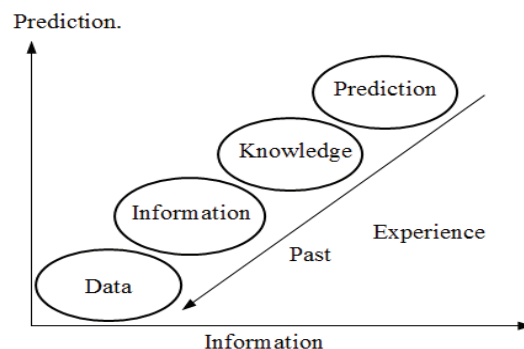

Fig. 1. Scheme of the transition from unstructured data to information and to knowledge

In Fig. 1. it is conditionally shown, that huge amount of data accumulates in time and is an experience, which is presented as a data. Such kind of the data allows to make a decisions about providing services with higher quality by telecom operator.

Relevant Mathematicals Methods for Big Data

To implement the collection processes, storage, processing, analysis and forecasting this Big Data processing methods are now widely used classification, clustering, neural networks, support vector machines, decision tree learning, others.

Machine learning (ML) methods are considered in many publications, for example, in [3, 4, 5, 6]. Essentially, this direction is intended to solve central task of intellectual system, which is evaluation of current object (situation).

The main expectation of ML methods is implementation of needs for flexible, adaptive, learning algorithms or calculation methods. As a result, new functions of systems and programs are provided.

The possibilities of ML methods, which are ability to learn and provide recommendations at the level of experts in a narrow subject area, provide algorithms, which are divided into two large groups:

- unsupervised learning (UL);
- supervised learning (SL).

Besides, sometimes such algorithms are identified:

- reinforcement learning (RL) [7];
- semi-supervised learning (SSL) [8].

The main task solved by the ML consists in assignment the observable object to one or another class for making a subsequent decision automatically or by a person. Such tasks are very widespread. An example is the analysis of user's actions when receiving services in electronic commerce systems, which allows making interface and the system's actions planning. In general, this analysis of data in various information systems allows making states predictions or objects classification.

UL methods solve the clustering task, when a set of previously not marked objects is divided into groups by an automatic procedure based on the properties of these objects. In this case the number of groups (clusters) can be previously defined or generated automatically. Such algorithms include the adaptive resonance theory (ART) and self-organizing maps (SOM) or Kohonen maps [9], and large group of clustering algorithms (k-means, mixture models, hierarchical clustering and etc.) [10, 11].

SL methods solve the classification task, when in potentially infinite aggregate of objects finite groups of designated objects are separated. Usually the groups' formation is performed by an expert. Thus, the expert can explain or can not explain for what reasons he has executed initial classification.

Using this initial classification as a model the classification algorithm must designate following unassigned objects as one or another expert-organized group based on the properties of these objects. SL methods include a large set of algorithms or algorithms families, which are often divided into linear and non-linear classifiers, that separate object classes depending on the form (hyperplane or hypersurface).

Approaches to the algorithms classification of ML methods are presented, in particular, in [12, 13].

In [14] the taxonomy of ML methods algorithms in the form of a hierarchical structure is considered, which does not pretend, of course, to be exhaustive completeness. Each of these algorithms essentially forms a family that can be modified for various needs of programs and algorithms that are often differed by computational complexity, implementation complexity and learning process automation, the ability to classify only two types (binary classification) or several objects types at once.

ML algorithms are costly from viewpoint of amount of computation when obtaining analytical and other results, especially when it is impossible to establish explicit envelopes and explicit relationships between the received data.

Evaluation practice of all components (technical, economic, social) influence is not widely accepted when analyzing development trends, service quality, etc. in terms of large amounts of information analysis.

This paper deals with the fuzzy logic approach that helps to reduce computational complexity in the process of classifying large amounts of data, the efficiency is considered on the example of obtaining a general estimation of providing services quality index.

To determine the quality index of providing services, it is assumed to obtain it based on the statistical information from the operator's monitoring system to minimize "human" factor influence. The received quality index of providing services determines the subsequent actions to improve the service quality based on the current value of the providing services quality index.

The providing services quality index can be used, for example, to estimate the required amount of cloud resources that are needed to connect, if there is a problem of the lack of resources of the existing infrastructure to satisfy the required providing services quality index.

To solve this problem, it is proposed to use decision-making methods based on fuzzy logic. Fuzzy expert rules are formed and they are the basis of the expert system.

Fuzzy rules in the fuzzy knowledge base can be periodically adjusted to the current status of the technical infrastructure by means of their reformation (refinement) in the process of providing services by telecom operator.

Because of the sufficient diverse analysis tasks facing the telecom operator, it is necessary to use an approach that will allow solving the described problems with the help of a certain set of mathematical methods and hence software tools developed on their basis.

In the proposed approach current status estimation of providing services by telecom operator is performed based on integral quality index and consists of the stages:

1. Quality index of providing services obtaining (Y).
   Summarized index characterizes the system status cumulatively (for example, service quality by telecom operator, economic indexes efficiency from the viewpoint of telecom operator and subscriber cumulatively) on the basis of measured private indexes (for example, using the operator's monitoring system).
   It is difficult for telecom operator to estimate the complex effect of individual quality indexes because it is necessary to know the weight coefficients or influence degree of each parameter on the general state of the communication service system cumulatively. Given this, integral quality index of providing services obtaining is a non-trivial task.

2. Fuzzy knowledge base (FKB) formation using the current values of integral quality index

   On the basis of the integral quality index, it is necessary to obtain an estimate of system status cumulatively and to understand the algorithms of its behavior, and also to compress information to knowledge or trends (based on logical or functional dependencies) that can be managed and made decisions. To compress large amounts of data it's possible to form FKB in the form of fuzzy logic rules. FKB uses fuzzy logic methods to obtain conclusions.

3. Estimation service quality by telecom operator using an integral quality index of providing services.

FKB formation is considered on the example of service quality estimation. Formation algorithm in the generalized form can be presented as follows:

Initial data: measurement tables provided by telecom operators. The measurement table is a set of parameters that can be presented by $X1 \ldots Xn$. It should be noted that $X1 \ldots Xn$ is set of parameters by which telecom operator estimates general state of the system, but none of them characterizes the system quality cumulatively (Y).

To develop FKB it is necessary to split measurement table into 3 samples:

1. Training sample with $M1$ rows, where $M1=\{1, k\}$, this data uses to form fuzzy logical rules of knowledge base. To form FKB rules it is necessary to determine integral quality index, which is obtained using desirability function denoted as YD;

2. Test sample with $M2$ rows, where $M2=\{k + 1, n\}$, which is needed to check fuzzy logical rules quality of knowledge base;

3. Examination sample with $M3$ rows, where $M3 = \{n + 1, m\}$, which is required for the final verification of the correct operation of the obtained FKB.

Generic algorithm for determining an unknown value of Y:

YD calculation using the desirability function based on $M1$ data.

FKB formation using $M2$, and obtained values of Y_{FKB}. FKB is formed by a set of rules:

«IF $X1, X2, X3, \ldots$, THEN Y ».

Therefor it is necessary to perform such steps:
- Clusterization.
- Membership function selection.

Y_{FKB} computation based on the selected membership function.

If $|YD - Y_{FKB}| > \varepsilon$, then transition to 2.2, where ε is acceptable deviation.

If $|YD - Y_{FKB}| \leq \varepsilon$, then the membership function was chosen correctly.

The main points of the proposed approach are described in more detail below.

I. Source Data

Such parameters of the monitoring system are used:
- *Connection Success Rate*, %
- *Connection Block Rate*, %
- *Connection Drop Rate*, %
- *PS Attach Success Rate*, %
- *PDP Context Activation Success Rate*, %
- *Speed DL*, kbitsps
- *Iub Congestion*, %
- *Backhaul Accessibility*, %
- *DNS Success Rate*, %
- *DNS Response Latency*, ms
- $W_{1\ldots10}$ – linguistic variables efficiency.

It is possible to form FKB for estimating the integral quality index and predicting trends of its change in a short time interval based on integral quality index parameters

Input information is the observation table $T = \{t_{M1}\}$, where i-th element $t_i = (x_i, y_i)$, $x_i = [x_{i1}, x_{i2}, \ldots, x_{ik}]$ — vector of input values, y_i — output value, M_1— number of observations, k — number of input variables, $x_i \in DX$, $y_i \in DY$.

They are needed to be converted to this values, when system can estimate service quality level cumulatively, because the data values in the table do not carry information. To calculate integral quality index of providing services it is proposed to use Harrington's desirability function.

Desirability Function

A specific feature of integral quality index determining task is that the operator does not know how to calculate Y. Such methods can be applied to determine Y:

1. Expert-weighted coefficients using (need to apply expert knowledge, which inject error probability due to human factor).

2. ML (the process of this method is smooth, but if there are emissions in the process, the algorithm becomes more complicated of such method).

The second approach is more flexible and allows getting the value of Y without using expert knowledge and allows reducing the error probability due to human factor.

To solve multicriterial tasks, various methods for constructing a generic index are used. One of them most

convenient methods is generic desirability function. It aroses as a result of observations of real solutions of experimenters and has such useful properties as continuity, monotonicity and smoothness.

The mathematical apparatus of converting specific parameters into abstract numerical values uses one of the logistic functions, which is called the "desirability curve":

$$d = exp[-exp(-Y)],$$

where coordinate axis Y is private indicators scale; axis d is desirability scale.

«Desirability curve» defines a function with two saturation sections (in $d{\rightarrow}0$ and $d{\rightarrow}1$) and linear section (from $d = 0.2$ to $d = 0.63$).

For evaluation integral quality index desirability scale is divided into five subranges in the range from 0 to 1:

[0; 0.2] – «*very bad*»,
[0.2; 0.37] – «*bad*»,
[0.37; 0.63] – «*satisfactorily*»,
[0.63; 0.8] – «*good*»,
[0.8; 1] – «*very good*».

The specific parameters of the compared systems are distributed on a scale corresponding to the requirements in the interval of effective values of the individual indexes scale. Then, the corresponding indexes are recalculated into grades on the desirability scale. The obtained value $d(i)$ for the i-th parameter is recalculated together with others into a generic desirability coefficient – D.

$$D = \sqrt{n \cdot d(1) \cdot d(2) \cdot \ldots \cdot d(n)} ,$$

where n is the number of used indexes.

The number of such indexes may not be the same for different systems. This allows comparing generic coefficients even when there are not some comparison parameters for different systems or data on them. The root of the n-th degree "smoothes" arise deviation. The obtained result allows evaluating systems (with a certain degree of accuracy).

Fuzzy Knowledge Base Formation

According to the peculiarities of Big Data processing, which is widely discussed in a number of works [15, 16], the approach based on fuzzy logic methods is proposed to consider the impact for explicit and fuzzy parameters and using the transition from unstructured data to FKB with explicit structured rules. This allows significantly reducing the amount of computing when data transmitting and data processing.

Initial data: fuzzy model is defined as a system with input variables $X = \{X_{k+1}, X_{k+2}, \ldots, X_n\}$, which are defined on the input area of discourses $DX = DX_{k+1} \cdot DX_{k+2} \cdot \ldots \cdot DX_n$, and one output variable Y, which is defined on the output area of discourses DY.

Explicit value is denoted as x_i for input variable X_i and as y for output variable Y.

A. *Clusterization*

There are many methods of clusterization. In this paper, a fuzzy c-averages algorithm is considered.

Fuzzy c-averages algorithm is based on distance minimizing from the observed data to cluster centers. For this, the Euclidean distance is calculated [17].

Fuzzy c-averages algorithm is as follows [17].

Given: observing table T, number of clusters c, number m, stop parameter ε.

Get: fuzzy partitioning matrix F and cluster centers matrix V.

Step 1. Fuzzy partitioning matrix F is initialized randomly. The fuzzy partitioning matrix F satisfies the conditions:

$$\forall_{k=1,\overline{M}}\left(\sum_{i=1}^{c}\mu_{ki}=1\right), \qquad \forall_{k=1,\overline{c}}\left(0<\sum_{i=1}^{c}\mu_{ki}<M\right)$$

Step 2. Calculate the cluster centers vector v_i.

Step 3. Calculate distance d_{ki} between k-th observation from the matrix T and i-th cluster center $d_{ki} = \sqrt{(t_k - v_i)}$.

Step 4. Calculate the matrix F approximation;

Step 5. If $\|F - F^*\|^2$, then the output from the algorithm, otherwise the transition to *Step 2*. Here F^* is fuzzy partitioning matrix, which is obtained at the previous step of the algorithm.

As a result of the algorithm cluster centers matrix $V = \{v_1, v_2, \ldots, v_c\}$ is calculated, where each $v_k = \{v_{k1}, v_{k2}, \ldots, v_{kn}, v_{kn+1}\}$, v_{ki} is coordinate value of i-th variable in k-th cluster, v_{kn+1} is coordinate value of output variable in k-th cluster. Each dedicated cluster corresponds to one linguistic rule fuzzy system rule base.

B. *Membership function selection*

Each linguistic term being in antecedent of the fuzzy rules is defined by its membership function.

There are several types of membership function: triangular, trapezoidal, piecewise linear, singleton, Gaussian.

As a part of the research Gaussian membership function was determined to be the best variant for estimation the integral quality index of providing services.

Based on the cluster analysis results, membership function parameters are defined as follows:

$$a=vij$$

$$\sigma_{ij}^2 = \frac{\sum\limits_{k=1}^{M}(\mu_{ki})^m(x_{kj}-y_{ij})^2}{\sum\limits_{k=1}^{M}(\mu_{ki})^m}$$

C. Y_{FKB} getting based on the selected membership function

Now, after determining the necessary parameters and membership functions constructing, the rules of fuzzy base are formed.

Fuzzy concept $LX_{i,k}$ corresponds to the judgment «X_i is approximately equal to v_{ij}». The general form of the rule is constructed in the form of a Mamdani type model.

R_j: $X_1 \approx v_{j1}$ AND … AND $X_n \approx v_{jn} \rightarrow y \approx v_{jn+1}$,

where v_{ji} is i-th input variable coordinate value in j-th cluster, v_{jn+1} is output variable coordinate value in j-th cluster, $1 \le j \le c$.

Thus, the proposed algorithm suggests solving the task of determining linguistic terms and forming a fuzzy rules base.

This approach in functioning process allows quickly determining which of parameters has made the most significant influence on the integral quality index. This allows managing technical aspects of network architecture. A similar approach can be applied to determine the economic and social parameters.

Fuzzy Knowledge Base Using for Service Quality Estimation

To obtain an integral quality index using FKB following conditions are necessary:

- there must be at least one rule for each linguistic term of the output variable
- for any term of the input variable there must be at least one rule, where this term is used as a antecedent;
- there should not be contradictions and correlations between rules.

The input variables are applied to the fuzzy logical inference algorithm input. These input variables carry information obtained in somehow, for example by measuring a certain physical quantity (operator's monitoring system). Resulting fuzzy logic output variables are formed at the algorithm output, for example satisfaction level of service quality. Thus, the fuzzy logic inference algorithm converts input variables into output variables based on certain fuzzy rules productions.

The main stages of obtaining fuzzy logical inference are the following:

1. Fuzzification (fuzziness introduction) is procedure, where for all input variables the membership functions specific values for each of linguistic terms are determined. Linguistic terms are used in in the set of rules base of fuzzy logical inference system.

2. Aggregation is procedure, when the truth degree of conditions for each rule of fuzzy logic inference algorithm. If the rule condition has a simple form, its truth is equal to the corresponding value of the membership function. If the condition consists of several subconditions of the form:

   RULE<#>: IF «b_1 is a_1» AND «b_2 is a_2», THEN «b_3 is V»,

   then truth degree is determined on the basis of known truth values of the subconditions and formulae for performing operations of fuzzy conjunction and fuzzy disjunction are applied:
   Fuzzy logic conjunction operation (*AND*):

   $$\mu_C(x) = \min\{\mu_A(x), \mu_B(x)\}$$

   Fuzzy logic c disjunction operation (*OR*):

   $$\mu_C(x) = \max\{\mu_A(x), \mu_B(x)\}$$

   For example, for a rule
   IF X1 = middle AND X2 = low AND X3 = low AND X4 = low
   The truth degree will be the following

   $$\mu_C(x) = \min\{0.5; 0.3; 0.4; 0.35\} = 0.3$$

3. Accumulation is the process of finding the membership function for each of the output linguistic variables.
   If conclusions relating to the same output linguistic variable belong to different rules of the fuzzy logic inference system, they are combined into fuzzy sets using the formula:

   $$y = \min\{x_m\},$$

   where x_m is modal value (moda) of Fuzzy set, which is corresponded to the output variable after the accumulation step, obtained in accordance with formula:

   $$x_m = \max\{\mu(x)\}, x \in [a, b]$$

   For example, for a rules:

   IF X1 = middle AND X2 = low AND X3 = low AND X4 = low, THEN Y = low $\mu(x) = 0.3$

   IF X1 = low AND X2 = high AND X3 = low AND X4 = high, THEN Y = middle $\mu(x) = 0.6$

   IF X1 = low AND X2 = middle AND X3 = middle AND X4 = high, THEN Y = middle $\mu(x) = 0.5$

Then $x_{m} = max\{0.3; 0.6; 0.5\} = 0.6.$

4. Defuzzification (leading to explicit) is a procedure for determining the usual (explicit) value for each of the output linguistic variables using different algorithms [18].
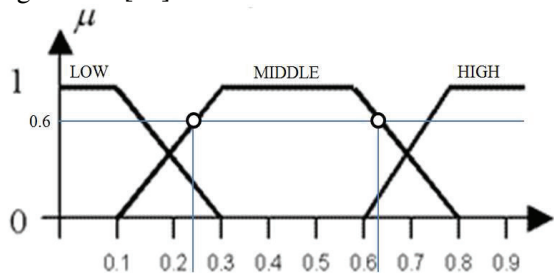

Fig. 2. An example of membership function

Fig. 2. shows, that the value of membership function can be corresponded to several arguments (if Y = middle from P.6, then at this stage only the membership function, which is responsible for the value of "MIDDLE" is used). There are many algorithms. In this paper right modal value method is used: $y = max\{x_i\}$.

Results

To verify the proposed approach, data from one of the Ukrainian telecom operators was used. Natural values were normalized in the range from 0 to 1. The fragment of the data is presented in Fig. 3.

X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
0,01	0,99	0,98	0,95	0,95	0,98	0,99	0,99	0,0075	0,99
0,001	0,9948	0,9989	0,9893	0,9958	0,9967	0,9967	0,9981	0,0033	0,9959
0,001	0,994	0,9989	0,9892	0,996	0,9972	0,9972	0,9968	0,0031	0,996
0,0006	0,9983	0,9993	0,9814	0,9643	1	0,9991	0,9992	0,0025	0,9984
...	...	...	...	...	...	...	...	...	...

Fig. 3. An example of initial data of Ukrainian telecom operator

According to the algorithm for determining unknown Y, the data were divided into three samples: teaching, test, examination.

From the first sample, Y was obtained using desirability function.

From the second sample fuzzy knowledge base was formed, which was consisted of the rules:

***IF** X1 = middle AND X2 = not low AND X3 = not low AND   X4 = not low   AND   X5 = not low   AND X6 = not low AND X7 = not low AND X8 = not low AND X9 = not very high AND X10 = not low,* **THEN Y = not low**

***IF** X1 = middle AND X2 = not low AND X3 = not low AND   X4 = not low   AND   X5 = not low   AND X6 = not low AND X7 = not low AND X8 = not low AND X9 = not very high AND X10 = not low,* **THEN Y = not low**

***IF** X1 = very low AND X2 = high AND X3 = high AND X4 = high AND X5 = high AND X6 = high AND X7 = high AND X8 = high AND X9 = very low AND X10 = high,* **THEN Y = high**

***IF** X1 = very low AND X2 = high AND X3 = high AND X4 = high AND X5 = high AND X6 = high AND X7 = high AND X8 = high AND X9 = very low AND X10 = high,* **THEN Y = high**

During the experiment, it was investigated that fuzzy knowledge base was formed for 4 iterations.

Correctness of the algorithm was tested by the third sample. The results of the algorithm fully corresponded to the expert estimates provided by one of the Ukrainian telecom operators.

Conclusions

As the researches result the approach for obtaining the integral quality index of comprehensive assessment of large amounts of structured data is proposed. The advantage of this approach is the information volume reduction processed during the Big Data analysis by bringing this amount of information to knowledge.

To conduct such an analysis time to set up and form a FKB is required, but this is compensated by a decrease of computational load during system functioning.

This approach considered allows reducing the computational complexity in the process of classifying large amounts of data. The desirability function using was suggested to solve the problem of obtaining the providing services integral quality index by the operator by a number of parameters that not be logically connected and that are provided by telecom operator's monitoring system.

Expert rules of the formed FKB have less influence of the human factor and more flexible react to rapid changes in the values of the certain service quality parameters in the process of system functioning.

Future work will be devoted to accounting economic and social parameters in conjunction with the technical characteristics of the telecom operator's infrastructure, and combination of three components (to solve the analytical problems that are important for telecom operator).

References

[1] Y.Y. Bolyubash, "Methods and tools for processing large data resources in information systems of territorial administration", Ph.D. dissertation, software for computer systems, Lviv Polytechnic, Lviv, Ukraine, 2017.

[2] M. Ulema. "Big Data and Telecommunications Telecom Analytics" – Tutorial II, BlackSeaCom'2016.

[3] S. Russell, P. Norvig, "Artificial Intelligence: A modern approach", New Jersey, Upper Saddle River, 2010.

[4] M. Tim Jones, "Artificial Intelligence: A Systems Approach", Hingham, Massachusetts, New Delhi, Infinity Sci. press LLC, 2008.

[5] D. Kriesel, "A Brief Introduction to Neural Networks", http://www.dkriesel.com/en/science/neural_networks.

[6] J. A. Cruz, Wishart D. S. Cancer Informatics, 2006 - 2:59–77.

[7] C. Szepesv`ari, "Algorithms for Reinforcement Learning", Synthesis Lectures on Artificial Intelligence and Machine. Learning series by Morgan & Claypool Publishers, 2009.

[8] X. Zhu, "Semi-Supervised Learning Literature Survey", Tech. Rep. 1530. University of Wisconsin-Madison, 2008.

[9] T. Kohonen, Biological Cybernetics - 1982 - 43(1):59–69.

[10] A.K. Jain, M. N. Murty, P. J. Flynn, "ACM Computing Surveys", - 1999 - 31(3):264–323.

[11] W.A. Barbakh, Y. Wu, C. Fyfe, "Studies in Computational Intelligence", - 2009 - 249:7–28.

[12] T.O. Ayodele, "Types of Machine Learning Algorithms", New Advances in Machine Learning. INTECH Open Access Pub., - 2010 - pp. 19–48.

[13] H.A.H. Ibrahim, S.M. Nor, A. Mohammed, A.B. Mohammed, International Journal of Computer Networks and Wireless Communications, - 2012 - 2(1):69–73.

[14] R.I. Mukhamediev, E.L. Mukhamedieva, Ya.I. Kuchin, « Taxonomy of machine learning methods and quality assessment of classification and learning », Electronic Journal "Cloud of Science", T. 2, № 3, 2015

[15] A.D. Mauro, M. Greco, M. Grimaldi, "What is Big Data? A Consensual Definition and a Review of Key Research Topics", 4th International Conference on Integrated Information, AIP Proceedings, 2014.

[16] B. Marr, J. Wiley, "Big Data: Using SMART Big Data, Analytics and Metrics To Make Better Decisions and Improve Performance", Sons Ltd, 2015.

[17] J. Espinosa, J. Vandewalle, V. Wertz, "Fuzzy logic, identification and predictive control", — London: Springer-Verlag, 2005. — 263 p.

[18] Y.A. Zak, "Alternative defuzzification Algorithms", System Research and Information Technologies. — 2015. — № 1. — pp.111-120. — Ref.: 6 titles — рос.

Глоба Л.С., Свецинська Є.С., Вольвач Є.О.

Розрахунок інтегрального показника якості надаваних сервісів

Проблематика. Дана стаття присвячена розгляду та удосконаленню існуючих алгоритмів оцінки якості надання сервісів телеком оператором. Сучасні підходи щодо визначення рівня якості послуг в системі телеком оператора потребують застосування складних математичних методів та підходів, які відрізняються обчислювальною складністю. У запропонованому підході оцінка якості надання сервісів базується на узагальненій функції якості, яка визначається за допомогою середнього геометричного індивідуальних показників якості системи моніторингу телеком оператора та дозволяє формувати нечітку базу знань у вигляді чітких правил. Для отримання значення інтегрального показника проводиться навчання системи правил нечіткої бази знань із застосуванням функції бажаності та методу кластерізації. Представлення набору даних, отриманого у визначені інтервали часу від системи моніторингу телеком оператора, у вигляді структурованих правил нечіткої бази знань дозволяє суттєво скоротити час прийняття рішення щодо стану якості надання сервісів, а також зменшити обчислювальну складність визначення якості послуг, які надаються кінцевому користувачу.

Мета досліджень. Підвищення якості надання послуг кінцевому користувачу за рахунок «м'якого» керування станом показників якості системи телеком оператора та зменшення обчислювальної складності під час визначення їх якості.

Методика реалізації. Дослідження виконано на основі аналізу великої кількості літературних джерел, теорії нечіткої логіки, методів кластерізації із застосуванням узагальненої функції якості, теорії баз нечітких знань.

Результати досліджень. Запропоновано підхід щодо визначення інтегрального показника якості послуг, які надаються кінцевому користувачу телеком оператором, можливості зведення комплексної оцінки значних обсягів погано структурованих даних до одного інтегрального нечіткого показника та формування на його основі системи знань, поданих у вигляді нечіткої бази знань. Під час формування нечіткої бази знань втрачається час на навчання її правил, але це компенсується зменшенням обчислювального навантаження на систему під час її функціонування.

Висновки. Проведене дослідження вказує на необхідність удосконалення реалізацій сучасної платформи надання послуг телеком оператором, яке полягає у введенні додаткового блоку «м'якого» керування їх якістю за допомогою нечіткої бази знань, яка дозволить замість обробки значних обсягів даних, отримуваних від системи моніторингу телеком оператора визначати якість послуг оперуючи правилами нечіткої бази знань.

Ключові слова: Big Data; база нечітких знань; нечітка логіка; якість обслуговування; інтегральний показник якості надання послуг

Глоба Л.С., Свецинская Е.С., Вольвач Е.А.
Расчет интегрированного показателя качества предоставляемых сервисов

Проблематика. Данная статья посвящена рассмотрению и усовершенствованию существующих алгоритмов оценки качества предоставления сервисов телеком оператором. Современные подходы к определению уровня качества услуг в системе телеком оператора требуют использования сложных математических методов и подходов, которые отличаются прежде всего вычислительной сложностью. Предлагаемый подход к оценке качества предоставляемых сервисов базируется на обощенной фукции качества, определяемой как среднее геометрическое индивидуальных показателей качества системы мониторинга телеком оператора и позволяет формировать нечеткую базу знаний в виде четких правил.

Для получения значения интегрального показателя проводится обучение системы правил нечеткой базы знаний с применением функции желательности и метода кластеризации. Представление набора данных, полученного в определенные интервалы времени от системы мониторинга телеком оператора в виде структурированных правил нечеткой базы знаний позволяет существенно сократить время принятия решения о качестве предоставления сервисов, а также уменьшить вычислительную сложность определения качества услуг, предоставляемых конечному пользователю.

Цель исследований. Повышение качества предоставления услуг конечному пользователю за счет «мягкого» управления состоянием показателей качества системы телеком оператора и уменьшения вычислительной сложности при определении их качества.

Методика реализации. Исследование выполнено на основе анализа большого количества литературных источников, теории нечеткой логики, методов кластеризации с применением обобщенной функции качества, теории баз нечетких знаний.

Результаты исследований. Предложен подход к определению интегрального показателя качества услуг, предоставляемых конечному пользователю телеком оператором, возможности сведения комплексной оценки значительных объемов плохо структурированных данных к одному интегральному нечеткому показателю и формированию на его основе системы знаний, представленных в виде нечеткой базы знаний. При формировании нечеткой базы знаний теряется время на обучение ее правил, но это компенсируется уменьшением вычислительной нагрузки на систему во время ее функционирования.

Выводы. Проведенное исследование указывает на необходимость совершенствования реализаций современной платформы предоставления услуг телеком оператором, которое заключается во введении дополнительного блока «мягкого» управления качеством услуг с помощью нечеткой базы знаний, которая позволит вместо обработки значительных объемов данных, получаемых от системы мониторинга телеком оператора определять качество услуг оперируя правилами нечеткой базы знаний.

Ключевые слова: Big Data; база нечетких знанний; нечеткая логика; качество обслуживания; интегральный показатель качества предоставления услуг.