System of Ontologies for Data Processing Applications Based on Implementation of Data Mining Techniques

Alexander Vodyaho[1], Nataly Zhukova[2]

[1]Saint-Petersburg State Electrotechnical University, Saint Petersburg, Russia
[2]Saint-Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, Saint Petersburg, Russia
{aivodyaho,nazhukova}@mail.ru

Abstract. The paper describes a system of ontologies developed for the applications oriented on solving problems of situations recognition and assessment based on results of data processing and analyses. Main attention is focused on the problems of processing measurements of various objects parameters represented in a form of time series. The considered applications process data using knowledge extracted from historical data with the help of Data Mining techniques. Such applications are highly knowledge centric and their core element is knowledge base that is represented as a system of ontologies. The proposed system of ontologies is a set of upper level ontologies for which techniques of adaptation for solving applied tasks for one or several related subject domains are developed.

Keywords: knowledge representation, data analyses, data fusion, measurements processing, situation recognition and assessment.

1 Introduction

Nowadays multiple problems in various subject domains are required to be solved at the level of situations [1, 2]. Results of solving problems at this level are much easier interpretable by an end user than results represented at lower levels of information generalization. Solving problems at the level of situations assumes solving such problems as recognition of situations, formal description of situations, analyses of situations, their estimation, assessment, prediction and awareness. Main sources of information about situations are results of measurements received from different types of instruments that measure parameters of technical and / or environmental objects. Real systems have to process huge volume of information including bad quality information. The majority of real life problems require that measurements are processed in real time or in the mode close to real time. It considerably increases the complexity of the problems. The problems can be solved with the desired quality and in limited time only using knowledge-oriented technologies. These intelligent technologies are based on application of data mining algorithms along with other means of artificial intelligence such as expert systems and inference machines. A set of basic solutions for developing intelligent technologies for measurements processing (IMPT) and examples of their implementation are proposed in [3, 4, 5, 6].

The intelligent measurements processing technologies are described in general form using web ontology language (OWL). When new measurements are received an appropriate technology is selected and detailed using an a priori defined set of production rules. The rules are two part structures that use first order logic for reasoning over knowledge representation [7]. The detailed technologies are processes described in business processes modeling language (BPML), they can be executed using standard engines. Execution of the processes requires that the input data, information and knowledge are represented using standard formats. It is reasonable to use the same standards for representing the results of measurements processing.

For formal description of data, information and knowledge about initial and processed measurements a hierarchy of information models has been developed [8]. In [6] a set of general classifiers for technologies, methods, algorithms and procedures for measurements processing is proposed. To use the intelligent technologies in the end user applications it is necessary to implement the models and to integrate them into the information models of the applications. For implanting the models it is proposed to use ontological approach as, at first, it has in fact become a standard for describing models of subject domains and, at second, the information models of the applications are commonly described using ontologies.

In the paper a structure of the system of ontologies build according to the models for measurements processing is proposed. Main data mining techniques and models required for measurements processing are enumerated in the second section. In the third section the developed system of ontologies is described. An example of the ontologies adaptation for the subject domain of telemetric information processing (TMI) is given in the fifth section.

2 Models and techniques for measurements processing and analyses

The actual standard of data and information processing and analyses is defined by the JDL model [9]. The JDL model is a general functional model of data and information fusion. The model has five levels: signal level, object level, situation level and level of threats. The highest fifth level is the level of decision making support. Measurements processing and analyses includes three steps: measurements harmonization, integration and fusion. Optionally measurements exploration can be executed at the fourth step. For each of the models levels, the functions and the processes of the levels are defined. The detailed descriptions of the models are given in [10] and the technologies of data harmonization, integration and fusion that provide the implementation of the models can be found in [11]. Input and output parameters of the levels of the functional models are represented using three specialized information models for description of different types of initial measurements and information and knowledge about them: a model of time series of measurements, a model of separate measurements and a combined model of different types of measurements. The description of each model is given in [3]. Processing of measurements at each level according to the developed technologies assumes application of an a priori defined set of intelligent technologies or separate statistical and data mining methods and algorithms adapted for solving tasks of measurements processing.

The set of intelligent technologies used for measurements harmonization is oriented on processing and analyses of initial binary data streams and the measurements represented in the form of single values or time series that are extracted from the streams. Processing and analyses of initial data streams assumes application of technologies for identification of the structures of the streams and estimation of the quality of the received data. Extracted measurements are transformed into standard formats and described in terms of the dictionary of the subject domain. Harmonization technology uses methods for measurements transformation into different formats, methods based on computing correlation functions, methods based on statistical laws of linguistic distribution, methods for building formalized descriptions of the initial data streams and measurements.

Intelligent technologies oriented on measurements integration include two key technologies: a technology for measurements preprocessing and a technology for preparing measurements for solving applied tasks. The first technology is implemented using algorithms of measurements denoising, removing single and group outliers, filling gaps, removing duplicating values and specialized procedures developed for different types of measurement instruments. The second technology uses methods for estimating compliance of the measurements to requirements of the end user tasks, methods for computing various features of measurements and characteristics of the analyzed objects.

Technologies of data fusion include technologies of extracting information and knowledge from initial measurements, of revealing dependencies in behavior of the measured objects parameters, of grouping measurements, of building grids on the base of separate measurements and of solving separate highly complicated computational tasks. The technology of extracting information and knowledge from measurements is based on algorithms of classification, cluster analyses and segmentation. The technology of revealing dependences applies algorithms of associations mining and building temporal patterns. The technology of measurements grouping is oriented on identifying groups of similar measurements and uses methods of cluster analyses. For the identified groups classes and association rules are defined. The technology of building grids is used to build both regular and non-regular hierarchical grids with various levels of detailing. The list of the computational tasks can include various tasks that are solved at the level of situations or oriented on decision making support.

The list of the technologies and methods given above is aimed to show the multiplicity of the directions of data mining techniques application for processing measurements. The detailed description of each technology one can find in [6]. The data, information and knowledge required to execute the methods and the algorithms directly affect the structure of the information models of measurements and results of their processing and, consequently, the structure of the system of ontologies for measurements processing.

3 A system of ontologies for measurements processing

The proposed interconnected ontologies are aimed to store and to provide data, information and knowledge about measurements and results of their processing. They are developed according to [12] and form the core of the system of ontologies of the subject domain of measurements processing. The system includes 3 main groups of ontologies: ontologies that contain information and knowledge about measurements, ontologies that describe technologies, methods, algorithms and procedures for measurements processing and analyses, and ontologies for representing information and knowledge about objects and situations using measurements of objects parameters. The first group contains the ontologies of time series, of time series segments, of time series features, of time series formal descriptions, of the criteria for the initial measurements and results of their processing estimation. The second group includes ontologies that provide information and knowledge about technologies of measurements processing, applied methods, algorithms and procedures including semantic descriptions of their input and output parameters, conditions of their application, the criteria for estimating results, the history of the methods application as well as other parameters. Ontologies of objects contain information about the structures of objects, their life cycles, functionality, possible interaction, defined regular states and faults. Ontologies of situations define the possible types of situations and provide extended formalized descriptions of situations and the objects involved in the situations.

Different kinds of external ontologies that are required for measurements processing or contain information about related subject domains can be used, for example, ontology of data providers or ontology of statistical distributions. For adaptation to applied subject domains the system can be extended with the specialized ontologies. The set of relations defined for the ontologies is given in Fig. 1.

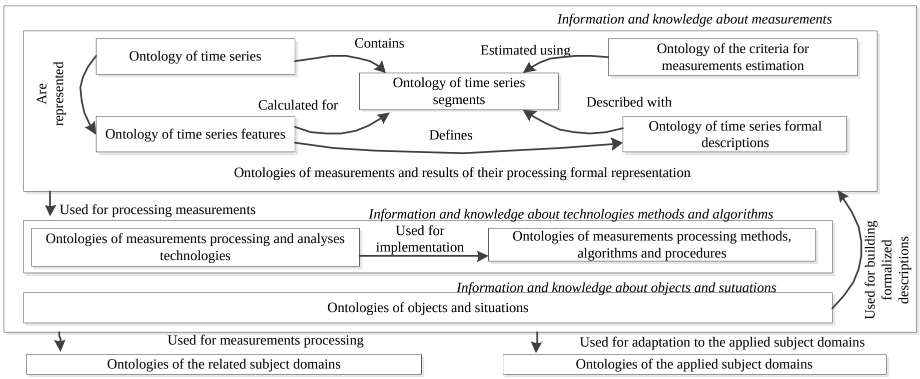


Fig 1. Relations defined for the system of ontologies

A. Description of the ontology of time series. The ontology of time series is aimed to provide information about different types of time series that can be processed. Types are formed according to behavior of time series and consequently define groups of algorithms that one can use for processing time series. The behavior of time series is described using five base features.

Feature 1. According to the types of the objects parameters 3 types of time series of measurements can be defined: functional, signal and constant. Functional time series are represented with continuous functions. For signal time series stepwise behavior is typical. Constant time series do not change in time.
Feature 2. Depending on dynamic of changes of functional time series slow changing time series and fast changing time series can be defined. The first type of time series can be characterized with the frequency spectrum in an interval from 0 up to 20-50 Hz, the second type – up to 2-3kHz or even more.
Feature 3. Depending on behavior, functional time series can be stationary, non-stationary and piece-wise stationary time series. The majority of time series are non-stationary but they contain comparatively long stationary segments.
Feature 4. For slow changing time series existence of gaps in the first and the second derivatives are considered as features.
Feature 5. For functional time series possibility of their description using parametric models is considered. For non-stationary time series a set of parametric models for each of the stationary segments is build. For selecting an appropriate model the models are matched using the least squares method or the method of maximum likelihood estimation.
For defining types of time series for each time series a set of various features is computed and classifiers of the time series types are used. The classifiers can be built on the base of historical data using algorithms for building decision trees [13].
B. Description of the ontology of time series segments. Segments are defined for piece-wise stationary and non-stationary time series. The ontology contains information about possible types of segments that can be observed in a time series. For defining types of segments 2 approaches are proposed. The first approach is based on using an a priori defined set of typical segments that are described in the ontology. To define a type of a segment, similar segments are found in the data base. The data base contains segments that have constant, linear increasing / decreasing, convexly / concavely increasing / decreasing behavior. The data base can be extended with segments that describe specialized behavior of time series typical for the applied subject domain. Specialized segments can be defined by experts or revealed from the historical data. The second approach assumes that for the analyzed segment a set of features is computed. The computed features contain several groups of features that reflect general behavior of the segment, describe the segment without taking into account the local peculiarities of the segment and that are focused on describing all tiny peculiarities of the segment. For defining methods and algorithms for computing features ontologies of methods are used.
C. Description of the ontology of time series features. The ontology is aimed to define features for describing stationary, piece-wise stationary and non-stationary functional time series and segments of time series. The sets of features computed for other types of time series, are fixed. The features can be defined according to the time required for features computing, according to the domain of the time series representation (time, frequency, time-frequency or spatio-temporal domain) and according to information density of the features for the solved task or for the allied subject domain.

The first group of features contains statistical features (median, mode, range, rank, standard deviation, coefficient of the variation, moments including mean, variance, skrewness, kurtosis), measurements frequency, behavior of the curve that corresponds to the time series in the time domain (convexity / concavity of the curve, variability of the curve, the error of the piece-wise constant / piece-wise linear approximation, the error of the approximation using the polynomials of the second and higher degrees, values of the characteristic points, the curvature), entropy, variability of the first derivative. The considered list of features contains commonly list feature, it can be extended or modified. The second group includes feature that consider time series as stochastic processes, in particular, one-dimensional and multi-dimensional distribution functions, one-dimensional and multi-dimensional probability density of the sophisticated processes, the distributions of the probabilities of the sophisticated discrete variables, spectral density. The list of features of the third group that are computed for both initial and transformed time series is given in table 1.

Table 1. Extended set of time series features

Transformation types	Computed features
initial measurements; ranging of values of initial measurements; computation of derivative using the finite difference method; computing of upper and lower envelopes	error of a time series description using a constant / linear / quadratic function for a time series approximation
computation of variation of upper and lower envelopes of a time series	deviation from zero
interpolation using cubic splines	error of interpolation transformation
approximation using a defined function; computation of a curve length	error of approximation transformation using power / exponential / logarithmic / user function
computation of a curve complexity	local complexity, global complexity and weighted complexity
computation of a curve variability	variability indices
computation of the characteristic points of a curve	number of minimums, maximums, intersections with the defined level of the values
computation of a curve curvature	minimum, maximum and median of a curvature
computation of area of a figure that is limited by the curve and the line that connects the edge points [14]	value of an area
computation of the first component using the method of principle components [15]	error of a time series description using a constant / linear / quadratic function for a time series approximation

The alternative approach for building the ontology of the time series features is proposed in [16]. It is based on computing linear, non-linear and other features. For defining linear features measures based on the computing of linear correlation, frequency parameters of the time series and autoregressive models are used. To nonlinear features refer 19 features. Definition of measures for these features assumes computation of nonlinear correlation and of time series dimension and complexity, building nonlinear models of time series.

D. Ontology of time series formal descriptions. The ontology is used for building formal descriptions of stationary, piece-wise stationary and non-stationary functional time series. Descriptions are built according to the computed features of the time series. The time series can be described using adaptive and non-adaptive approaches [17]. Adaptive approach assumes computing coefficients of piece-wise constant and piece-wise linear approximations, coefficients of singular decomposition and building symbolic representations of time series. In order to build non-adaptive descriptions one can use such features as coefficients of wavelet transformations, of time series spectral representation, results of piece-wise aggregate approximation. Depending of time series complexity one or several descriptions can be built.

E. Description of the ontology of criteria for initial measurements and results of their processing estimation. In the ontology 3 groups of criteria for initial measurements are considered. The first group allows one to estimate measurements using knowledge about the object / environmental area which parameters are measured, the second group − using results of matching new data with historical data, the third group − using specialized procedures selected according to the types of the processed measurements and applied methods. The criteria of the first group are usually defined by experts and / or producers of the measurement instruments. They are represented as a set of features for which admissible intervals for measured values are given. The second group of criteria is based on computing distances between the analyzed measurements or their features and measurements that were acquired earlier in similar conditions. The third group of the criteria includes criteria that estimate separate measurements and sets of measurements, separate time series and their groups. The criteria significantly depend on the solved tasks. The examples of the criteria are uniqueness, accuracy, consistency, completeness, timeliness, actuality, interpretability, relatedness to other data.

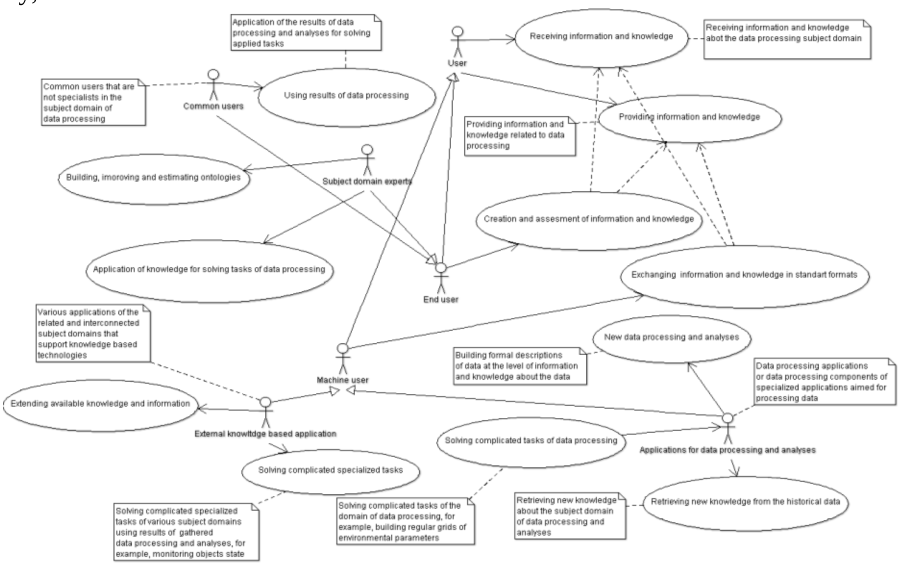


Fig 2. Use case diagram for the system of the ontologies for measurements processing

Results of measurements processing are estimated twice: just after measurements are processed and at consequent stages of their processing and analyses. Both stages assumes application of the procedures of revealing contradictions of the acquired results

with available information, of comparing results received using different methods, of comparing results with results of historical data processing, of comparing results of separate measurements and separate time series processing with the results of joint analyses, of computing complex features on the base of separate features. An example of criteria for cluster analyses methods can be found in [18].
The described above system of ontologies but can be used for solving tasks in intelligent applications specialized for measurements processing by experts and common users and by different external applications. The use case diagram for the proposed system of ontologies is given in Fig. 2.

4 Application of the system of ontologies for TMI processing

The developed set of ontologies for measurements processing was adapted for processing TMI [19] received from remote space objects. A hierarchy of the solved tasks is given in Fig. 3.

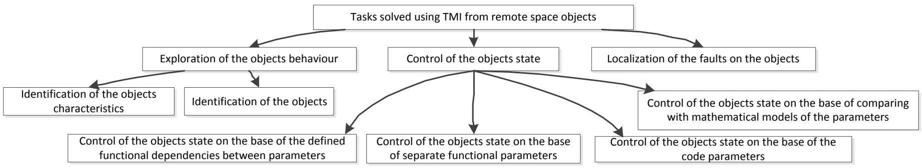


Fig 3. Ontology of the tasks

Table 2. Time series of measurements of specialized parameters

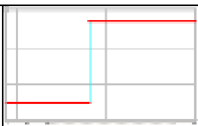
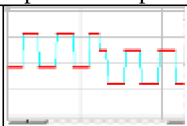
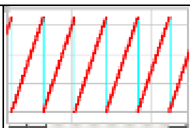
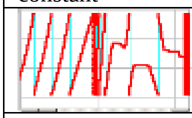
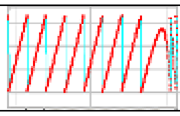
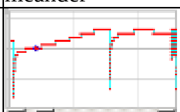
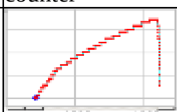
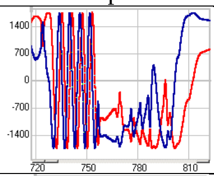
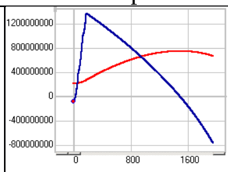
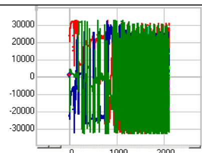
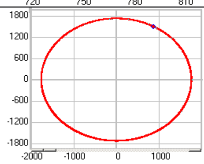
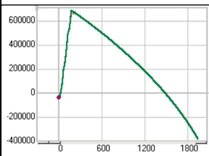
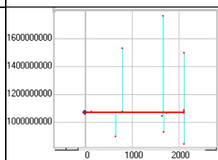
			
constant	code	meander	counter
			
mantissa	order	lower part	upper part

Table 3. Standard dependences of telemetric parameters

Example of the initial data			
Dependency graphical re-presentation			
De-pen-dency	pairs of sine and cosine $x^2 + y^2 = 1$	integro-differential pairs $x' - y = 0$	elements of the matrix of the coordinates transformation

Adaptation required extension of the ontology of the types of times series, the ontology for representing dependences in objects parameters and the ontology of methods and algorithms for measurements processing. A set of types of time series was extended with the types aimed to describe measurements of specialized parameters (table 2). The set of features for the specialized types are defined in [20]. The standard dependencies of telemetric parameters include pairs of sine and cosine, the integro-differential pairs and elements of the matrix of the coordinates transformation (table 3). The upper level ontology of methods and algorithms for TMI processing is given in Fig.4. Several branches of the ontology are detailed in Fig. 5-7.

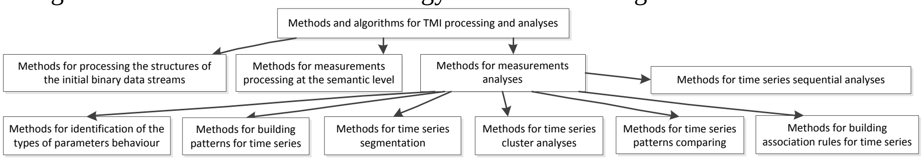


Fig 4. Ontology of methods and algorithms for TMI processing and analyses

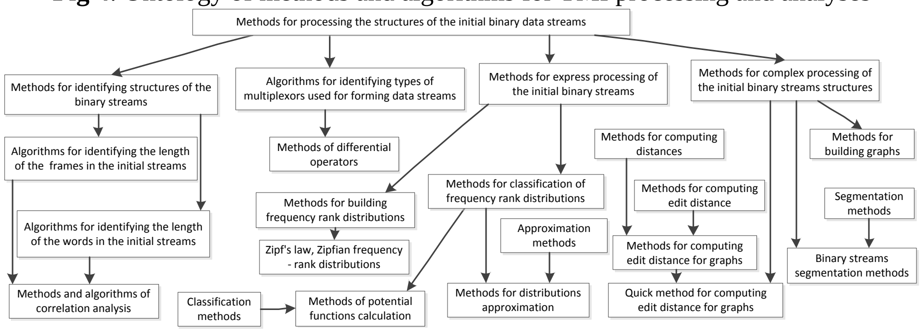


Fig 5. A fragment of the ontology of methods for processing structures of binary streams

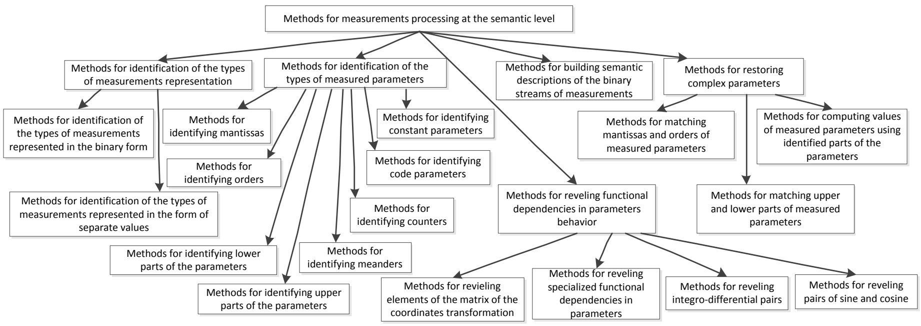


Fig 6. A fragment of the ontology of methods for measurements processing at the semantic level

The system of the ontologies was implemented in a number of the applications oriented on processing TMI from space objects in the delayed mode that are successfully used for about ten years already. The description of the developed systems and the examples of their application can be found in [6, 21].

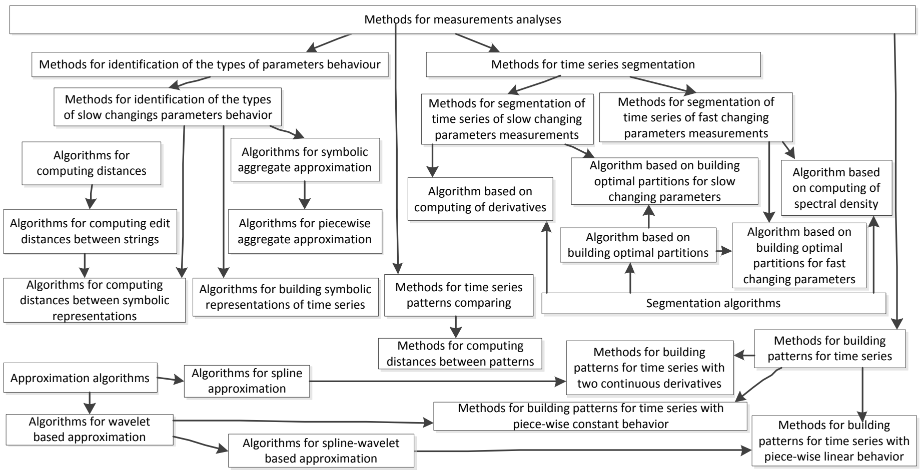


Fig 7. A fragment of the ontology of methods for measurements analyses

5 Case Study

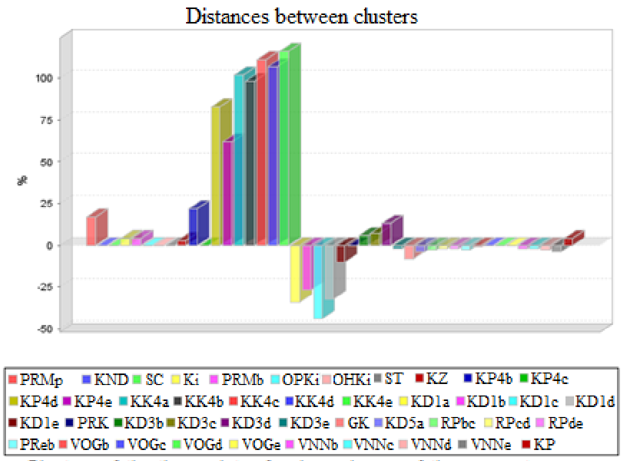
The control of the space objects state using code parameters assumes analyses of the time points at which the values of the parameters changed. These points correspond to the moments of execution of commands on the controlled objects. In table 4 a subset of code parameters for three different objects of one type are given. For each parameter the time points of their values change are defined.

Table 4. The time of the values change points of the code parameters

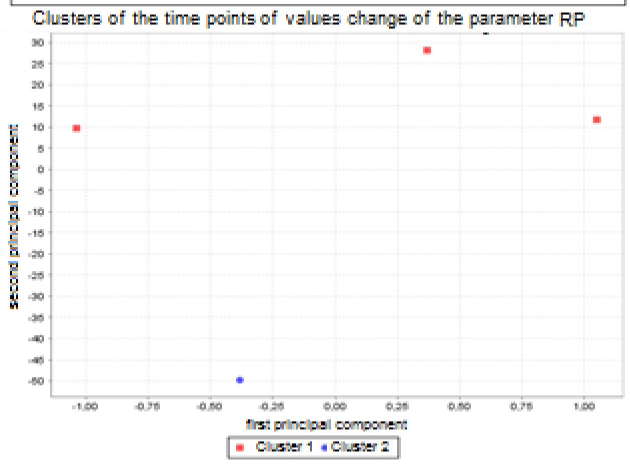
№	*PRMp*	*KND*	*SC*	*Ki*	*PRMb*	*OPKi*	*OHKi*	*ST*	*KZ*
1	362789	344936	348956	350428	350429	359539	359535	361435	361746
2	453563	464518	468542	470111	470113	479124	479121	481018	481328
3	190444	201398	205418	206915	206917	216025	216018	217898	218208
	KP4b	*KP4c*	*KP4d*	*KP4e*	*KK4a*	*KK4b*	*KK4c*	*KK4d*	*KK4e*
1	361479	361478	361483	361483	361944	361943	361940	361944	361955
2	481061	481061	481062	481063	481475	481476	481476	481477	481477
3	217941	217941	217942	217943	218355	218356	218357	218357	218358
	KD1a	*KD1b*	*KD1c*	*KD1d*	*KD1e*	*PRK*	*KD3b*	*KD3c*	*KD3d*
1	362327	362373	362366	362387	362388	362789	363040	363040	363042
2	481930	482010	481990	481991	481970	482372	482623	482605	482606
3	218832	218833	218870	218891	218871	219252	219482	219497	219482
	KD3e	*GK*	*KD5a*	*RPbc*	*RPcd*	*RPde*	*RPeb*	*VOGb*	*VOGc*
1	363042	363042	363295	363300	363299	363307	363300	363330	363329
2	482645	482653	482906	482910	482908	482915	482909	482933	482927
3	219494	219504	219746	219748	219746	219755	219750	219777	219775
	VOGd	*VOGe*	*VNNb*	*VNNc*	*VNNd*	*VNNe*	*KP*	-	-
1	363330	363330	363332	363331	363332	363330	363356	-	-
2	482919	482930	482940	482938	482938	482939	482950	-	-
3	219778	219777	219781	219780	219784	219783	219791	-	-

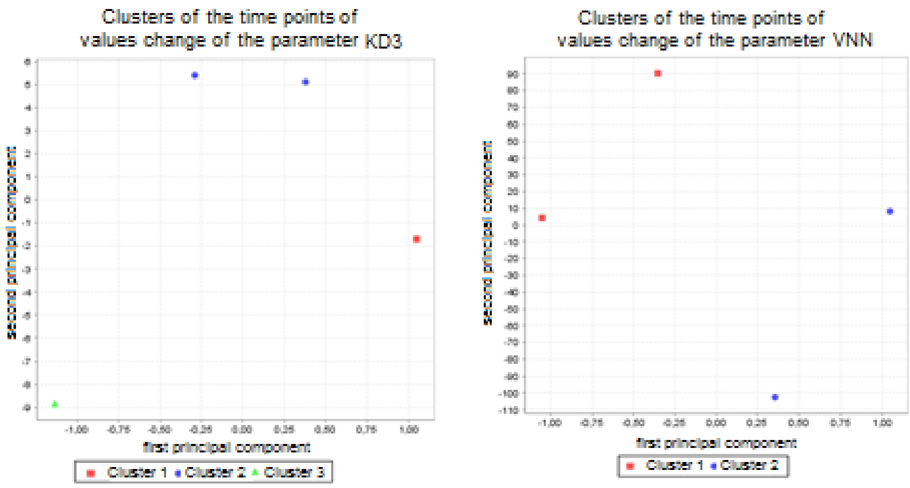
The time points of the values change were processed using data mining techniques, in particular, statistical and cluster analyses methods. The results of building clusters of objects using all parameters showed that the behavior of the first object differs significantly from the behavior of the second and the third objects. The first object is the only element of the first cluster. The second and the third objects form the second cluster. The differences between the clusters are represented in the form of a histogram (Fig. 8 a). The order of the parameters in the histogram is the same as in the table 4. The cluster analyses of similar parameters of different blocks of the objects that have equal construction (the name of the block to which the parameters refer is written in small letters after the name of the parameter) revealed deviations from the normal behavior for the parameters RPde (the time points of the disconnection of the spherical locks between blocks ‘b’ and ‘e’ differ from the time points defined for the same parameter between other blocks), KD3 (the time points of the contacts breaking of blocks ‘b’ and ‘d’ differ from the time points defined for the parameter for blocks ‘c’ and ‘e’), VNN (the time points of the output of the tooth for blocks ‘d’ and ‘e’ differ from the time points defined for blocks ‘b’ and ‘e’) (Fig. 8 b-d). The clusters in Fig. 8 are represented in the feature space build using the principal component method [22].

a)



b)





c) d)

Fig 8. Application of Data Mining techniques for processing time points of the values change of code parameters

6 Conclusion

In the paper a system of ontologies required for processing and analyzes of various objects parameters measurements represented in the form of time series or single values is presented. The structure of the ontologies and the relations between the ontologies that link them into a system are defined. For each of the ontologies a detailed description is provided and the relations with external ontologies are enumerated.

The proposed system of the ontologies has the following distinguishing features:

- the system allows one to solve the tasks of measurements processing taking into account the peculiarities of the processed data and the solved tasks;
- multiple technological solutions for measurements processing based on application of intelligent methods and algorithms can be implemented using the considered set of ontologies;
- the structure of the system of the ontologies and of the separate ontologies is simple and can be easily extended and modified if new methods are developed or new types of measurements are defined;
- information and knowledge represented in the form of ontologies can be interpreted both by experts and machines and can be multiply used;
- the system of ontologies can be easily adapted to different subject domains if ontological descriptions of the domains are available.

Further development of the described system of ontologies assumes detailing the ontologies on the base of knowledge, acquired as a result of operating of the developed applications for telemetric information processing. A set of applications for other subject domains is going to be developed and approved.

References

1. Steinberg A.N. Foundations of Situation and Threat Assessment, Handbook of Multisensor Data Fusion, D. Hall, M. Liggins, J. Llinas (eds.), LLC Books (2008).
2. Steinberg, A.N. ; Rogova, G. Situation and context in data fusion and natural language understanding. Proceedings of 11th FUSION, Cologne (2008).
3. Vitol A., Zhukova N., Pankin A. Adaptive multidimensional measurements processing using IGIS technologies. Proceedings of the 6th International Workshop on Information Fusion and Geographic Information Systems: Environmental and Urban Challenges, St. Petersburg (2013)
4. Pankin A., Vodyaho A., Zhukova N. Operative Measurements Analyses in Situation Early Recognition Tasks. Proceedings of the 11th International Conference on Pattern Recognition and Image Analyses, Samara (2013)
5. Zhukova N. Method for adaptive multidimentional meas-urements processing based on IGIS technologies. Proceedings of the 11th International Conference on Pattern Recognition and Image Analyses, .Samara (2013)
6. Vitol A., Deripaska A., Zhukova N., Sokolov I. Technology of adaptive measurements processing. SPbSTU «LETI», Saint-Petersburg (2012)
7. Browne P. JBoss Drools Business Rules. Packt Publishing (2009)
8. Vitol A., Zhukova N., Pankin A. Model for knowledge representation of multidimensional measurements processing results in the environment of intelligent GIS. Proceedings of the 20th International Conference on Conceptual Structures for Knowledge Representation for STEM Research and Education, Mumbai (2013)
9. Steinberg A., Bowman C., White F. Revisions to the JDL Data Fusion Model. Sensor Fusion: Architectures, Algorithms, and Applications. Proceedings of the SPIE, vol. 3719 (1999)
10. Zhukova N. Harmonization, integration and fusion of multidimensional measurements of technical and natural objects parameters in monitoring systems [in Russian]. Izvestiya SPbETU "LETI", vol 2, Saint-Petersburg (2013)
11. Popovich V., Voronin M. Data Harmonization, Integration and Fusion: three sources and three major components of Geoinformation Technologies. Proceedings of IF&GIS, St. Petersburg (2005)
12. http://www.w3.org/
13. Quinlan R. C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, San Mateo (1993)
14. Feng S., Kogan I., Krim H. Classification of curves in 2D and 3D via affine integral signatures. Acta Applicandae Mathematicae, vol 109, issue 3, Springer, Nitherlands (2010)
15. Chang K., Ghosh J. Principal curve classifier - a nonlinear approach to pattern classification. Proceedings of Neural Networks, Anchorage (1998)
16. Kugiumtzis D., Tsimpiris A. Measures of Analysis of Time Series (MATS): A MATLAB Toolkit for Computation of Multiple Measures on Time Series Data Bases. Journal of Statistical Software, vol. 33, issue 5 (2010)
17. Lin J, Keogh E., Lonardi S., Chiu B. A Symbolic Representation of Time Series, with Implications for Streaming Algorithms. Proceedings of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery, San Diego (2003)
18. Halkidi M., Batistakis Y., Vazirgiannis M. Clustering Validity Checking Methods. ACM Sigmod Record 31(2,3) (2001)
19. Nazarov A., Kozyrev G., Shitov I. et al.: Modern Telemetry in Theory and in Practice. Training Course [in Russian]. Nauka i Tekhnika, St. Petersburg (2007)

20. Vasiljev A., Vitol A, Zhukova N. Detecting the symantic structure of the group telemetric signal [in Russian]. SPbSTU «LETI», Saint-Petersburg (2010)
21. Vasiljev A., Geppener V.,Zhukova N.,Tristanov A.,Ecalo A. Automatic control system of complex dynamic objects state on the base of telemetering information analysis [in Russian]. 8th International Conference on Pattern Recognition and Image Analysis: New Information Technologies, vol.2, No.4 (2007)
22. Jolliffe I. Principal Component Analysis. Springer, 2nd ed. (2002)

Система онтологий для приложений обработки данных на основе техник анализа данных

Александр Водяхо[1], Наталья Жукова[2]

[1] Санкт-Петербургский государственный электротехнический университет, Санкт-Петербург, Россия
[2] Санкт-Петербургский институт информатики и автоматизации Российской академии наук, Санкт-Петербург, Россия
{aivodyaho,nazhukova}@mail.ru

Аннотация. В статье описана система онтологий, спроектированных для приложений, ориентированных на решение проблем распознавания и оценки ситуаций на основе результатов обработки и анализа данных. Основное внимание сосредоточено на проблемах обработки измерений от различных объектов с параметрами, представленными в виде временных рядов. Рассмотренные приложения обрабатывают данные при помощи знаний, извлечённых из исторических данных при помощи техник анализа данных. Такие приложения очень зависят от базы знаний, представляющей собой систему онтологий. Представленная система онтологий является множеством онтологий верхнего уровня, для которых разработаны способы решения задач в одной или нескольких предметных областях.

Ключевые слова: представление знаний, анализ данных, слияние данных, обработка измерений, распознавание и оценка ситуаций.