Use of Scan Statistics in Intelligent Heating Control Systems

Aleksejs Jurenoks, Svetlana Jurenoka

***Abstract*—**"Intelligent heating system" have widely established their position as a research field during the last decade. Nowadays the technical solutions related to energy resource management are being rapidly developed and integrated into the daily lives of people. The methodology to provide fast online fault detection in intelligent heating control system data-set using scan statistics method is described. Scan statistics have long been used to detect statistically significant bursts of events. This research system work scenario faults in time enables to detect the most problematics scenarios of temperature calibration process, check the efficiency of the decisions taken to select the work strategy and proposes to use a context sensitive and proactive fuzzy control system for controlling temperature actuator in heating system using.

***Keywords*—** Energy Efficiency, Fuzzy logic, Sensor network.

I. Introduction

The number of studies related to intellectual and adaptive system use in the smart home context is increasing daily. Nowadays it is not difficult to imagine a situation when a person could live in a house where all processes are managed by artificial intelligence. Most of the time the smart system of house management system will be able to react correctly to the actions of the human or take proper proactive actions, both based on predesigned model and measured variables.

Scan statistics are a powerful method for detecting unusually high rates of events, also called anomalies [1]. Scanning for bursts of events has many applications in diverse fields such as telecommunications, astronomy, quality control, reliability and e-learning [7]. In this paper described method use scan statistics to monitor the occurrence of events in time, a such as status messages, alarms, and faults. As result system can automatically identify unusual bursts in events during educational process.

Fuzzy set theory provides a major newer paradigm in modeling and reasoning with uncertainty. Though there were several forerunners in science and philosophy, in particular in the areas of multivalued logics and vague concepts, Lotfi A. Zadeh, a professor at University of California at Berkeley was the first to propose a theory of fuzzy sets and an associated logic, namely fuzzy logic [9]. Essentially, a fuzzy set is a set whose members may have degrees of membership between 0 and 1, as opposed to classical sets where each element must have either 0 or 1 as the membership degree—if 0, the element is completely outside the set; if 1, the element is completely in the set. As classical logic is based on classical set theory, fuzzy logic is based on fuzzy set theory.

This article describes the fuzzy control method for heating circuit management, which can adapt to the user's needs by acquiring new operating requirements. In order to use the system, scan statistics method of data selection are used; it can add or remove requirements or modify the existing ones by using the information received from the environment in real time.

II. Related works

The first industrial studies related to fuzzy logic using fuzzy control systems were initiated in the 1980s when two Danish engineers L.P. Holmblad and J.J. Østergaard developed a fuzzy controller for automated cement kilns with the option to control the specific burning temperature during the production process. The results were first published in 1982 [3]. The results did not obtain a great interest in Europe, but gained a lot of popularity in Japan where the developed methodology was improved and fuzzy logic was applied in the automation process of the Sendai City subway train management. The developed product gained a lot of support and the technology that was applied was adapted for other similar management systems which were based on the classic On – Off operating principle [4]. This success contributed to the great interest by the industry representatives regarding the feasibility of introducing the Fuzzy logic into the control system automation processes and as a result they were used in industrial control systems as well as in facility heating/ventilation systems.

In the 1990s, the Japanese started to integrate the fuzzy controller in everyday products like washing machines, camcorders, vacuum cleaners as well as transport; as a result, in 1992, a successful technology implementation experience in Japan increased the interest of Europe and the US regarding the fuzzy controller technology and this lead to the fuzzy logic implementation in Europe, in 1993 in the area of information systems.

When designing automation systems more and more attention has been paid to the Fuzzy system [2, 3, 9] because these systems are well suited for the environment where the dynamic values can change unexpectedly and the next step cannot be predicted or calculated mathematically. Many articles have proven that Fuzzy systems are quite easy to understand and design because they are based on natural language. Control systems that use fuzzy logic in the automation process are generally fast, user friendly and cheap because they do not need much memory and resources to operate. [8]

III. Application of Scan Statistics Method

Let us introduce N events, distributed at time interval (0, T). In this paper *Sw* is defined as maximal number of events at a time interval with length *w* (the window of fixed length *w* of time). The maximum set *Sw* is called the scan statistics where one scan process is done in one period of time (0, T) with a window of size w and observes a large number of points. *Wk* is the shortest period of time containing a fixed number of *k* events. The distributions of the statistics *Sw* and *Wk* are related. If the shortest window that contains *k* points is longer than *w*, then there is no window of length *w* that contains *k* or more points:

$$\Pi(\Omega_\kappa > \omega) = \Pi(\Sigma_\omega < \kappa). \tag{1}$$

Let us introduce the distribution of intelligent heating system working condition faults using data received from that was implemented in private house sector in Latvian period from 01.09.2016 till 31.12.2016. In reviewing the data note that there is a 1 period (from 12.12.16 through 16.12.16) when 80 faults were registered (see Figure 1). It is justified by the fact, that in this period of time almost in all household temperature calibration process was done.

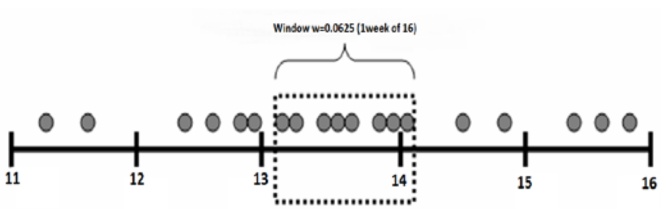

Fig. 1. Scanning the unit time interval with the window of length w=0.0.625. Black points represent times of occurrence of N=190 events, S0.0625=80. The centers of the occurred “points” Ci have coordinates t1, t2, …, tN

In Figure 1 we can see concentration of 8 * 10 points at the time interval from 12.12.16 till 16.12.16. There is a question whether it is possible to explain such concentration proceeding from a null hypothesis [1]. If it is impossible to explain the given concentration of points (faults) by means of a null hypothesis it is necessary to recognize that the given concentration of points (faults) has a special character. It means that process of occurrence of failures in the given situations is influenced by additional factors.

The given conclusion is an objective signal for decision-making in the sphere of data flow management in e-learning system. We might explain this as follows: each of the 190 cases could either fall in the period from 01.09.2016 to 31.12.2016 or not, independently of the other cases. The probability b (k, N, w) found by computing the binomial probability for N = 190, p = 1/16:

$$b(k,N,w)=\binom{N}{k}w^k(1-w)^{N-k}= \\ b(k,190,0.2)=\binom{19}{k}0.2^k(1-0.2)^{19-k} \tag{2}$$

It is easy to understand that there is an limitation number of sliding windows during the time interval. To solve this problem in a constructive way we must define some limitation set of sliding windows (Figure 2).

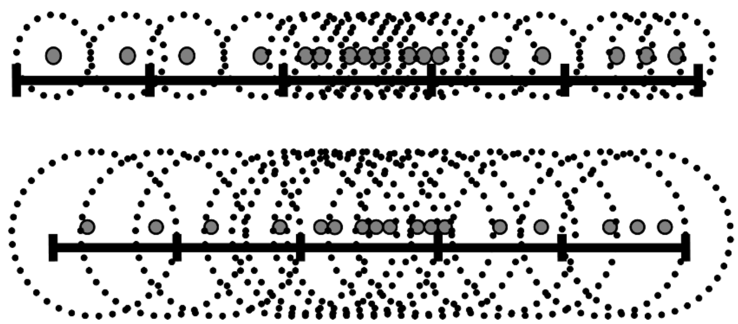

Fig. 2. Illustration of the scanning window of two fixed lengths w=0.0,312 and w=0.0625. The centers of scanning windows are points with time coordinates t1, t2, …, tN

Let us introduce P(k, 190, 0.0625) the distribution of the maximum number of cases in one period of time. The null hypothesis model for 190 cases C_i (i=1, 2,…, 190) were distributed independently by the binomial distribution function and completely at random over the 4 month period. If the cluster is k=80, the total number of cases over the whole 4 month period is N=190 and the window size w is 1 week out of 16, or w=0.0625. The result provided by the probability P (80, 190, 0.0625) can be calculated using formula 3:

$$P(k,N,w)\approx(N-k+1)b(k-1,N,w) \\ -(N-k-1)b(k,N,w)+2G_b(k+1,N,w) \tag{3}$$

where

$$b(k,N,w)=\binom{N}{k}w^k(1-w)^{N-k}$$

$$G_b(k,N,w)=\sum_{i=k}^{N}b(i,N,w)$$

Analyzing the statistics of system faults, it is possible to get an exact analytical solution for the distribution of “incorrect / problematics” questions in database. In this case there exists only one possibility and it is to use intelligent module to make modelling in real time. In the paper we have considered to use the Monte-Carlo method of scan statistics for calculation of p-value and testing null hypothesis H_0 (no clusters). The authors the approach by Wallenstein and Naus in assuming the null hypothesis model which can be used for investigations of similar problems.

To isolate situations causing problems it is necessary to analyze the level of unsolved tasks in e-learning system using method of scan statistics. The paper illustrates the scanning process with a circle window with fixed radius. The illustration of the Monte-Carlo algorithm for cluster detection is presented in Figure 3.

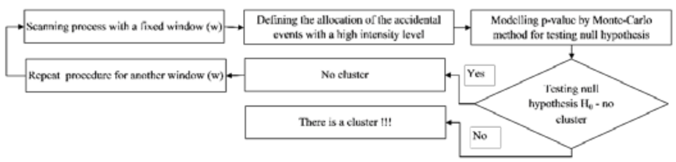

Fig. 3. Illustration of the Monte-Carlo algorithm for cluster detection

Using scan statistics method with heating system [5] knowledge database intelligent module will be able to classify operating scenario priority tags. In such way undefined or fault scenarios will be tagged to lover metric. Such approach will help to:

- to detect objects in knowledge database with utmost fault / unsolved intensity;
- to check significance of fault detection (user process interruption or cancelation) with highest frequency;
- to analyze the dynamics of changes of clusters detected taking into consideration the time factor.

IV. CONTROLER CALLIBRATION METHOD

This article describes the heating boiler control system, which uses three linguistic values relating to space: low, adequate, high and two linguistic values relating to water temperature: hot, cold. Let us look at a set of requirements that uses the input data of the linguistic variables. As a result, the management conditions of the distribution valve by using water temperature and the room temperature are as follows.

```
IF (room temperature = “high” && water temperature =
“hot”) THEN close control valve.

IF (room temperature = “low” && water temperature =
“hot”) THEN open control valve.

IF (room temperature = “adequate”) THEN do nothing.

IF (room temperature = “high” && water temperature =
“cold”) THEN close control valve.

IF (room temperature = “low” && water temperature =
“cold”) THEN close control valve.
```

Referring to an expert opinion, the affiliation of the physical values of the temperature will be defined to the set of fuzzy values: “low” = 0,1; “adequate” = 0,5; “high” = 0,9; “cold” = 0,1; “hot” = 0,9 (figure 4.)

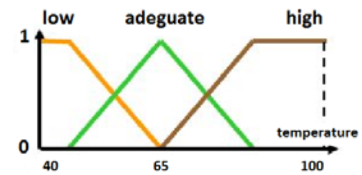

Fig. 4. Heating system valve control linguistic variable description

Using the defined values, the conditional terms will be transformed using rules:

```
Rule 1 : Close control valve min(0,9;0,9);
Rule 2 : Open control valve min(0,1;0,9);
Rule 3 : Do nothing: 0,5 = 0,5;
Rule 4 : Close control valve min(0,9;0,1);
Rule 5 : Close control valve min(0,1;0,1);
```

A valve servo motor ESBE ARA 600 was used in the experiment that ensured the maximum turning step which is equal to 90 degrees. For the actuator to work, 120 seconds are required in order to execute a full cycle. As a result, the procedure “close control valve” means rotating the valve by 90 degrees and the procedure “open control valve” means rotating the valve by -90 degrees. Using the industrially defined analogue management system during the experiments has proven that the system provides inconsistent temperature changes which do not allow stabilizing the temperature of the heating system using simple if condition to control servo motor (Figure 5). This is based on the engine performance time (120 sec.) and the time of thermal inertia in the heating system.

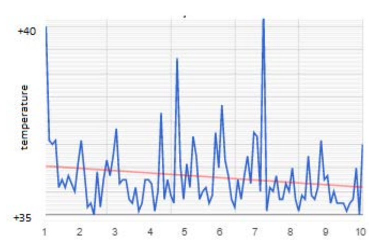

Fig. 5. Temperature calibration using simple logical rules decreasing heating temperature of heating boiler.

To avoid wide temperature variations, the optimal valve rotation angle for every system iteration should be determined by using the formula 4.

$$\alpha = \frac{average\ value * \max value + 0.1 * range_n}{average\ value + \min value} \tag{4}$$

As a result, the data of the experiment have determined that the first rotation step of the iteration = 15,75 degrees. During the experiment, the control system can operate only with integers; this results in the first iteration rotation step being equal to 16 degrees and it is necessary for the execution of the iteration for 22 seconds. Consequently, the remaining operating range of the valve = 74 degrees and the time required for an action is 98 sec.

By completing each iteration cycle, standby time t_{wait} is required and it provides accurate reading of the information by taking into account the inertia change of the heating system temperature using formula 5.

$$t_{wait} = \frac{t_1+t_2+\cdots+t_n}{n}, where\ t_{wait} > t_n - t_{n-1} \quad (5)$$

where $t1$ is the time between the temperature changes.

In an operating system new physical values of the temperature are defined for the fuzzy variables. Table 1 shows the operating principle of the current method which chooses the optimal valve operating range using the pre-defined conditions by reducing the temperature fluctuations.

TABLE I. VALVE CALIBRATION STEPS.

Cold	Hot	Adequate	Operating range	Operating step
0,1	0,9	0,5	90	16
0,1	0,75	0,425	75	15
0,15	0,75	0,45	60	11
0,15	0,75	0,45	60	11
0,3	0,75	0,525	45	6
0,3	0,65	0,475	35	5
0,38	0,65	0,515	27	4
0,38	0,59	0,485	21	3
0,42	0,59	0,505	17	3
0,42	0,55	0,485	13	2
0,5	0,55	0,525	5	1
0,5	0,55	0,525	5	1
0,5	0,55	0,525	5	1
0,5	0,55	0,525	5	1

It is known [6, 8] that the thermometric temperature sensors are not able to instantly intercept the temperature changes of the particular environment as the heat exchange between the environment, the thermometric surface happens with finite rate and it takes time for the temperature sensor to read the correct temperature value. Suppose that the temperature sensor with the temperature T_0 is placed in a new environment with a temperature of Θ. As a result, the sensor temperature will change by using the formula 6:

$$T - \theta = (T_0 - \theta)e^{-\frac{\tau}{\lambda}} \quad (6)$$

where Θ – the real temperature of the environment, τ – time, λ – temperature sensor thermal inertia factor.

In order to determine the temperature sensor inertia factor, the activation time of the system is (t = 0). A certain temperature difference has been determined, if the initial temperature difference T_0 and θ is known.

Consequently, the adaptive system is able to determine the inertia changes of the temperature by using the Formula 7.

$$n = \frac{T_0-\theta}{T-\theta} = e^{\frac{\tau}{\lambda}} \quad (7)$$

As a result, using the method for determining the operational temperature range as well as the inertia factor of the temperature sensor the automation control system was designed.

V. ADAPTIVE SUBSYSTEM EXPLOITATION IN AUTONOMOUS ADAPTIVE CONTROL SYSTEMS.

For an adaptive control system to function, the training cycle and the operational calibration of the starting system for certain tasks and the environment – it is necessary to define a set of requirements and values of the starting system which contains general knowledge of the object and the operating environment. During the initial stage a system is able to meet the basic management principles at the beginning of the life cycle of a system or in case of emergency by using the sensor nodes. It is known that the management of a system is not effective when using only the predefined constant values in comparison with a system that works with the help of the empirical knowledge basis.

In order to create the control system of the initial phase, the article describes a situation where poorly formalized management objects are used. The main task of the system is to provide minimum object management possibilities at the early stage of the system life cycle; as a result, a control system which uses fuzzy logic principles was applied.

The system input data are the peak operating angle is l of the management engine and the maximum full time of the rotation cycle is t_{max}.

The basis of the control module consists of the fuzzy variables F1..Fn, dF1..dF2 described in this article characterizing the boiler temperature, room temperature and the linguistic values of the valve position:

$$F_N \ \&\ dFM \rightarrow AN*M \quad (6)$$

where N and M are unclear sets that contain information about the certain room temperature and boiler temperature value range.

In order to select an executable requirement A from the requirement sets the following formula is used:

$$A = \frac{\sum(A_k \cdot P_k)}{\sum P_k} \quad (7)$$

where Ak – the number of an executable requirement, Pk – the output values of the current requirement.

In cases when the system expert wants to limit the range of decision, the system is assigned with the requirement having the maximum value which will result in the completion of the task that matches the requirement.

VI. Controller Practical Validation

To test the methodology described in this article a boiler control module is developed in which the control algorithm was implemented by using the Arduino language. In order to implement all the requirements of the algorithm ArduinoUno microcontroller and the analogue temperature sensor LM335 were used; to visualize the operational processes LCD displays (1602B) were used. The actuator was implemented by using ESBO ARA 600 servomotor which was used to regulate the heating valve.

A. The time of the experiment

The first stage – September 2016, with the average outdoor air temperature + 13,2 C^o. The heating system was in the standby mode most of the time and the highest consumption of the fuel was used to maintain the boiler combustion process (Fig 6.)

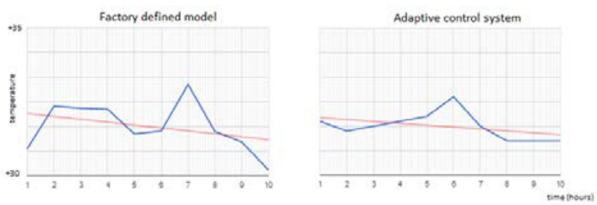

Fig. 6. The first phase of the temperature variation in the heating system.

The second stage – November 2016, with the average outdoor air temperature + 0,6 C^o. The heating system operates using 50% of its nominal capacity (Fig.7).

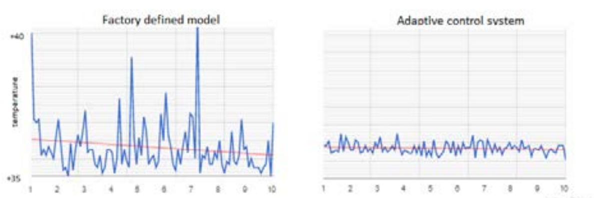

Fig. 7. The second phase of the temperature variation in the heating system.

The third stage – 7th January, 2017. The average daily air temperature –18,2 C^o. The heating system operates using 90% of the nominal capacity (Fig.8).

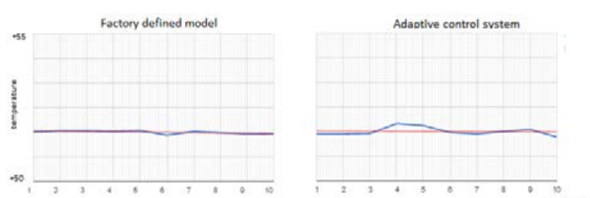

Fig. 8. The third phase of the temperature variation in the heating system

B. Benchmark value determination

The time t_{et} (min) will be assumed as the reference value – how long the heating system can operate using the Q quantity of fuel. Let us assume that the heating system operates via industrially defined ON-OFF valve control algorithm and the system can be considered as active as long as the room temperature stays within the given border.

As a result, each stage has certain benchmark values of the system performance, which are also taken from the boiler output temperature fluctuations.

TABLE II. BURNING BOILER FUEL CONSUMPTION

$T_{et1\ (min)}$	$Q_{1\ (kg)}$	kg/min	$T_{et2\ (min)}$	$Q_{2\ (kg)}$	kg/min	$T_{et3\ (min)}$	$Q_{3\ (kg)}$	kg/min
260	16	0,06	133	16	0,12	82	16	0,20
502	32	0,06	250	32	0,13	158	32	0,20
758	48	0,06	378	48	0,13	228	48	0,21
958	64	0,07	508	64	0,13	300	64	0,21
1218	80	0,07	592	80	0,14	398	80	0,20
1470	96	0,07	718	96	0,13	462	96	0,21
1712	112	0,07	873	112	0,13	515	112	0,22

C. The summary of the experiment results

By using the control algorithm of the adaptive heating system mixer described in this article, the overall increase in the system performance expectancy in relation to the benchmark model (Table 3) can be seen, but there are cases when the heating system is operating at a nominal capacity and the mixer valve is opened to 90%; then by using the adaptive valve control algorithm with the calibration tasks the excess energy is consumed.

TABLE III. ADAPTIVE MANAGEMENT SYSTEM APPLICATION RESULTS.

$Q_{(kg)}$	$t_{et1\ (min)}$	$t_{fuzzy1\ (min)}$	$t_{et2\ (min)}$	$t_{fuzzy2\ (min)}$	$t_{et3\ (min)}$	$t_{fuzzy3\ (min)}$
16	260	↑ 264	133	↑ 148	82	↑ 83
32	502	↑ 510	250	↑ 268	158	↓ 142
48	758	↑ 762	378	↑ 392	228	↓ 219
64	958	↑ 962	508	↑ 535	300	↓ 298
80	1218	↑ 1224	592	↑ 618	398	↓ 390
96	1470	↑ 1478	718	↑ 738	462	↓ 462
112	1712	↑ 1722	873	↑ 896	515	↓ 505

VII. Conclusions

This article has described and practically tested the fuzzy control algorithm for the management of the heating system mixer module for adaptive management that can be adapted to the user's needs while acquiring new operating rules. In order to use the system, it does not need prior training, the training process takes place during the operation of the system by modifying the pre-defined requirements using the information received real-time via the sensors. The article proves that the effectiveness of the control algorithm depends on the boiler operating temperature and the outside air temperature. In cases when the boiler operates using nominal energy, the control of the operational process increases the consumption of fuel resources wasting the heat energy through heating temperature calibration.

The long-term use of the adaptive control module of a heating system, in systems operating on 40 to 60% of the rated power is obtained by decreasing the consumption of the heating fuel by 8% (~1920 kg/year) by using a 50 kW heating system.

The data obtained will be used for further research by developing intellectual management methodology of the universal heating system.

ACKNOWLEDGMENT

This work has been supported by the European Regional Development Fund within the Activity 1.1.1.2 “Post-doctoral Research Aid” of the Specific Aid Objective 1.1.1 “To increase the research and innovative capacity of scientific institutions of Latvia and the ability to attract external financing, investing in human resources and infrastructure” of the Operational Programme “Growth and Employment” (No. 1.1.1.2/VIAA/1/16/043)

REFERENCES

[1] Berman, M. and. Eagleson, G. K. Dec., 1985. A Useful Upper Bound for the Tail Probabilities of the Scan Statistics. Journal of the American Statistical Association, Vol. 80, No. 392, doi: 10.2307/2288548, 886-889.
[2] Collier, S.E. The Emerging Internet: convergence of the smart grid with the Internet of Things // Rural Electric Power Conference (REPC), 2015 IEEE, pp. 65–68
[3] Holmblad L.P.. J.J. Østergaard In Fuzzy Information and Decision Processes (M. M. Gupta and E. Sanchez, Eds)., pp. 389 – 399. North-Holland, New York, 1982
[4] Halder S., Ghosal A., Chaudhuri A., DasBit S. A probability density function for energy-balanced lifetime-enhancing node deployment in WSN // Proceedings of the 2011 International Conference on Computational Science and Its Applications, ICCSA’11. Berlin, Heidelberg: Springer-Verlag, 2011, pp. 472–487.
[5] Jurenoks, A., Novickis, L., Jurenoka, S. Modelling the performance of adaptive multi-boiler heating system. In: *Sistemas e Tecnologias de Informação*, Portugal, Lisabona, 21-24 June, 2017. Universidade do Minho, Lisabona: 2017, pp.1553-1559. ISBN 978-989-98434-7-9.
[6] Jurenoks A., Novickis L. Wireless sensor network live circle simulation tools to balance energy consumption in the network nodes // Preceedings of 29th European Simulation and Modelling Conference (ESM 2015), Leicester, Great Britan, 26–28 October 2015, pp. 245 – 249. ISBN 978-9077381-908.
[7] Jurenoka, S., Jurenoks, A., Novickis, L. Online fault detection methodology from Moodle quiz system using scan statistics method. No: Information and Software Technologies, Druskininkai, 12.-14. oktobris, 2017. Tiergartenstraße 17, 69121 Heidelberg, Germany: Springer, 2017, 478.-486.lpp. ISBN 978-3-319-67642-5.
[8] Mautz R., Tilch S. Survey of optical indoor positioning systems // International Conference on Indoor Positioning and Indoor Navigation (IPIN), 2011,pp. 1–7.
[9] Mäntyjärvi, J., Seppänen, T. Adapting applications in handheld devices using fuzzy context information, Interacting with Computers, vol. 15, issue 4, p. 521-538, August 2003

Dr. Aleksejs Jurenoks is an assistant professor at Riga Technical University faculty of Computer science and Information technology. He obtained Dr.sc.ing. degree in 2016. His research interests include sensor network, adaptive and intelligent systems of information processing and applied systems. He is the author of 28 publications and 6 educational books.
He is regularly involved in different EU-funded projects: eLOGMAR-M (FP6, 2004 -2006); IST4Balt (FP6, 2004 -2007), UNITE (FP6, 2006 - 2008) BONITA (INTERREG, 2008-2012). Contact information: phone: +371 29606250, E-mail: aleksejs.jurenoks@rtu.lv

Mg.sc.ing. Svetlana Jurenoka is a lector at Riga Technical University faculty of Computer science and Information technology. She is PhD student of computer science since 2017. Here research interests include adaptive and intelligent systems of information processing and e-Learning systems. She is the author of 4 publications.
He is regularly involved in different EU-funded projects: eLOGMAR-M (FP6, 2004 -2006); UNITE (FP6, 2006 - 2008); BONITA (INTERREG, 2008-2012). Contact information: phone: +371 29525703, E-mail: svetlana.jurenoka@gmail.com