Preprints, 12th IFAC International Workshop on
Adaptation and Learning in Control and Signal Processing,
June 29 - July 1, 2016. Eindhoven, The Netherlands

A Dynamic Threshold Based Algorithm for Change Detection in Autonomous Systems

Alexander Kalmuk * **Oleg Granichin** * **Olga Granichina** ** **Mingyue Ding** ***

* *Saint Petersburg State University (Faculty of Mathematics and Mechanics, and Research Laboratory for Analysis and Modeling of Social Processes), 7-9, Universitetskaya Nab., St. Petersburg, 199034, Russia. Oleg Granichin is also with Institute of Problems of Mechanical Engineering (IPME RAS), St. Petersburg, Russia (e-mails: alexkalmuk@gmail.com, o.granichin@spbu.ru)*
** *Herzen State Pedagogical University, St. Petersburg, Russia (e-mails: olga_granichina@mail.ru)*
*** *Huazhong University of Science and Technology, Wuhan, Hubei 430074, China (e-mails: myding@hust.edu.cn)*

Abstract: The paper deals with the detection of abrupt changes in dynamical systems in the presence of external noise. The system is considered as a black box. That is, it is supposed that nothing is known about a structure and relations within the system, and we can just obtain the values of some parameters of the system. A new algorithm is proposed in the paper to deal with such highly uncertainty. At first, we use moving average with adaptive window size, then GLR (Generalized Likelihood Ratio) is used to detect abrupt changes in the obtained data using an adaptive threshold. This approach is applied to slowdown detection of a small autonomous car with only accelerometer on the board.

Keywords: Uncertain systems; Fault detection; Estimation

1. INTRODUCTION

The problem of change detection is to determine changes in the characteristics of a dynamical system during time. Often, such characteristics can be variation in mean value and variance of the distribution of some parameters of the system. This study is close in meaning to fault detection. But in case of fault detection a fault occurs only inside the system, while in case of change detection there are no difference what is a source of change.

The problem of abrupt changes detection is highly important, because an abrupt change may be a result of some underlying or external problem. This problem is well studied, and many approaches are proposed (see Akimov and Matasov (2015), Basseville and Nikiforov (1998), Blanke et al. (2006)). Some of them suppose that probability density function depends upon a scalar parameter, while others suppose multidimensional probability density function. In this paper we consider the first one - a case of scalar parameter.

In this paper it is assumed that the dynamical system is considered as a black box. That is, we do not know how the parameters of the system are related to each other, and we can only measure some of these parameters. Also it is assumed that measurable data contains Gaussian noise, which is a practical and reasonable assumption. In this conditions it is impossible to apply such well known approaches as Kalman filter, because the system is just a black box. Therefore, one of the possible way could be to use moving average to filter noise firstly. For this purpose we use spatial adaptive estimation of nonparametric regression (see Goldenshluger and Nemirovski (1997), Lepskii (1990)), which proposes how to select adaptive window size. After we got filtered data, we are interested in the two following things - a method for detection an abrupt change and selection of a threshold. There are many ways to detect an abrupt change on the filtered data (see Basseville and Nikiforov (1998)). We have selected GLR, because it shows good properties when the actual value of a parameter after change is unknown (it is supposed that it is from some predefined set). Finally, after we detect a change on the filtered observations, a question about threshold in the presence of external noise arises. That is, some constant threshold is used in GLR, but we remember that there is a noise initially. So we need to choose a new dynamic threshold, such that it would be no worse than constant one in the presence of noise.

The idea of reconstruction of a constant threshold to the dynamic one is not new. Such thresholds can be used to reduce the delays associated with the constant threshold method (see Perhinschi et al. (2006), Verdier et al. (2008), Shu et al. (2008)) and for more accurate estimation in the presence of noise (see Davis et al. (2006)).

The main contribution of the paper lies in the combining of Generalized Likelihood Ratio algorithm with an idea of dynamic threshold obtained by moving average with

[1] This work was supported by Russian Science Foundation (project 16-19-00057).

Copyright © 2016 IFAC

adaptive window size to create more accurate and efficient method for abrupt changes detection.

We have made an experiment to show how this approach works in the real world. We consider a small autonomous car, that detects its own slowdown. The car is based on STM32F3-Discovery with accelerometer inside. It moves along a smooth road and at some point faces with a piece of a foam rubber. It does not stop, but slows down. We detect this moment of time and stop the motor.

The article is organized as follows. Section 2 presents the problem statement. In sections 3 and 4 an overview of used methods is provided. Section 5 presents the main result. The results of numerical experiments are shown in Section 6, followed by Conclusions and Future work discussion.

2. PROBLEM STATEMENT

Let us consider the dynamical system with the model of observations. Because we consider the system as a black box, we can only deal with measurable part. So let us assume that we have a sequence of independent random variables Y_t, which can be measured at each time instance.

Our goal is to determine a fault in the observable variables Y_t. Each of variables $Y^i = (y_1, y_2, ..., y_m)$ have the probability density function p^i_θ depending upon only one scalar parameter $\theta \in \mathbb{R}$. Before the unknown moment of time t_0, the parameter θ is equal to θ_0. The problem is to detect a change of the parameter θ.

We solve this problem in two steps:

(1) Filter a noise out of Y_t.
(2) Apply Generalized Likelihood Ratio to the obtained data.

3. SPATIAL ADAPTIVE ESTIMATION OF NONPARAMETRIC REGRESSION

Let us consider the following problem. We want to restore a signal function from noisy observations. Suppose there are noisy observations $y(x)$ of a signal function $f(x)\colon [0,1] \to \mathbb{R}$ – along the regular grid $\Gamma_n = \{i/n, i = 0, ..., n\}$:

$$y(x) = f(x) + \xi(x), \tag{1}$$

where $\{\xi(x)\}_{x\in\Gamma_n}$ is a sequence of independent $\mathcal{N}(0,1)$ random variables defined on the underlying probability space (Ω, A, P).

In Goldenshluger and Nemirovski (1997) the estimates by the least square method of $f(x)$ at a given point $u \in [0,1]$ are considered. Suppose that the degree of estimate is 1. So we get the following approximation at a given point x_0:

$$\hat{f}_\Delta(x_0) = \frac{1}{N_\Delta} \sum_{x\in M_\Delta} y(x), \tag{2}$$

where $\Delta \in [0,1]$ is some segment $[x_0 - \delta, x_0 + \delta]$ centered at x_0 and containing at least one observation point, M_Δ is the set of observation points in Δ, N_Δ is the cardinality of M_Δ.

The problem is to select "the best" window when no a priori information on f is available. Let us introduce the following estimation:

$$|\hat{f}(x_0) - f(x_0)| \le \omega_f(x_0, \delta) + N_\Delta^{-1/2}|\zeta(\Delta)|, \tag{3}$$

where $\zeta(\Delta) = \frac{1}{N_\Delta}\sum_{x\in M_\Delta} \xi(x)$, $\omega_f(x,\delta) = \sup_{x\in\Delta} |f(x) - f(x_0)|$. The right hand side is comprised of two terms – deterministic $\omega_f(x,\delta)$ and stochastic error $N_\Delta^{-1/2}|\zeta(\Delta)|$. Since $\zeta(\Delta)$ is $\mathcal{N}(0,1)$, the stochastic error typically is of order of $(n\delta)^{-1/2}$:

$$P\{N_\Delta^{-1/2}|\zeta(\Delta)| > \kappa(n\delta)^{-1/2}\} \le exp\{-c\kappa^2\}, \tag{4}$$

with certain absolute constant $c > 0$. Now, there are no more than n essentially different (resulting in different sets M_Δ) choices of Δ. Let these choices be

$$\Delta_1 \subset \Delta_2 \subset ... \subset \Delta_N,$$

and let $2\delta_1, 2\delta_2, ..., 2\delta_N$ be the length of the windows $\Delta_1, \Delta_2, ..., \Delta_N$. We obtain

$$\Omega_\kappa = \{\omega \in \Omega \mid N_{\Delta_i}^{-1/2}|\zeta(\Delta_i)| \le \kappa(n\delta_i)^{-1/2}, \quad where\ i = 1, ..., N\}. \tag{5}$$

Assuming that $\omega \in \Omega_k$, (3) can be strengthen as

$$|\hat{f}_{\Delta_i}(x_0) - f(x_0)| \le \omega_f(x_0, \delta_i) + \kappa(n\delta_i)^{-1/2}, \tag{6}$$

Notice that as i grows, then the first term in the right hand side increases, and the second term decreases; therefore a reasonable choice of the window to be used is that one which balances both the terms, say, the one related to the largest i with $\omega_f(x_0, \delta_i) \le \kappa(n\delta_i)^{-1/2}$.

4. GENERALIZED LIKELIHOOD RATIO

This method is based on the likelihood ratio test (see Granichin et al. (2015)). The main reason to use it is that the parameter θ_1 is unknown after change. Let us introduce log-likelihood ratio for the observations from time j up to time k is

$$S_k^j = \sum_{i=j}^{k} \ln \frac{p_{\theta_1}(y_i)}{p_{\theta_0}(y_i)} \tag{7}$$

In the present case, θ_1 is unknown; therefore, this ratio is a function of two unknown independent parameters : the change time and the value of the parameter after change. The standard statistical approach is to use the maximum likelihood estimates of these two parameters, and thus the double maximization:

$$g_k = max_{1\le j\le k} \sup_{|\theta_1-\theta_0|\ge\nu>0} S_k^j(\theta_1) \tag{8}$$

and the following stopping rule:

$$t_a = \min\{k : \sum_{i=0}^{N-1} I_{\{g_{k-i}\ge h\}} \ge \eta\}, \tag{9}$$

where I is an indicator function, h is a threshold for the derivative, and η is a threshold for the number of crossings of h, and ν is a known minimum magnitude.

5. THE ESTIMATE OF A SIGNAL

In this section we will introduce a dynamic threshold and show how it depends on a constant threshold in the presence of standard Gaussian noise. To do this we apply (6) for the estimation f of the input signal (2). And then we will use this estimation in Generalized Likelihood Ratio.

Theorem 1 in Goldenshluger and Nemirovski (1997) says how to estimate $|\hat{f}(x_0) - f(x_0)|$. Now we will reformulate this theorem in the case when the degree of estimation is 1 (i.e. (2)).

At first, let us define the ideal window $\Delta_k^*(x_0) = [x_0 - \delta_k^*, x_0 + \delta_k^*]$ for the new conditions - $p = l = 1$, $\theta_m = 1$:

$$\delta_\kappa^* = \max\Big\{\delta \le \delta_0 : \frac{\kappa}{\sqrt{2n\delta}} \ge 4\int_{x_0-\delta}^{x_0+\delta} |f(x)|dx\Big\}. \quad (10)$$

One can check that coefficients $\alpha_\Delta(x,u)$ (see 3.1 Construction in Goldenshluger and Nemirovski (1997)) will be equal to $\frac{1}{N_\Delta}$. That means that the index of the window Δ defined as follows

$$r_{\Delta,u} = \Big(\sum_{x\in M_\Delta} \alpha_\Delta^2(x,u)\Big)^{1/2} = \frac{1}{\sqrt{N_\Delta}}. \quad (11)$$

So Ξ will be defined as follows:

$$\Xi = \Big\{\omega \in \Omega : \frac{1}{\sqrt{N}}|\sum_{x\in M_\Delta} \xi(x)| \le \kappa\Big\}. \quad (12)$$

In this case we get the following estimation using Theorem 1 in Goldenshluger and Nemirovski (1997) with conditions (10), (11), (12):

$$|\hat{f}(x_0) - f(x_0)| \le \frac{6\kappa}{\sqrt{2n\delta_\kappa^*(x_0)}}, \ when\ \omega \in \Xi. \quad (13)$$

Now we want to set up an adaptive threshold to apply it in Filtered Derivative Algorithm described above. Assuming that $\hat{f}$ has Gaussian distribution $\mathcal{N}(\mu,\sigma)$, we can write out new expression for g_k:

$$\begin{aligned} g_k &= \max_{1\le j\le k} \sup_{|\mu_1-\mu_0|\ge\nu>0} \sum_{i=j}^{k} \ln\frac{p_{\theta_1}(y_i)}{p_{\theta_0}(y_i)} \\ &= \max_{1\le j\le k} \sup_{|\mu_1-\mu_0|\ge\nu>0} \sum_{i=j}^{k} \ln\frac{\frac{1}{\sigma\sqrt{2\pi}}\exp^{\frac{(y_i-\mu_1)^2}{2\sigma^2}}}{\frac{1}{\sigma\sqrt{2\pi}}\exp^{\frac{(y_i-\mu_0)^2}{2\sigma^2}}} \\ &= \max_{1\le j\le k} \sup_{|\mu_1-\mu_0|\ge\nu>0} \sum_{i=j}^{k} \frac{y_i(\mu_0-\mu_1)}{\sigma^2} + k\frac{\mu_1^2-\mu_0^2}{2\sigma^2}. \end{aligned}$$

where μ_0 and μ_1 are the mean values of the distribution corresponding to probability density functions p_{θ_0} and p_{θ_1}.

Let us denote $S := \frac{\mu_0^*-\mu_1^*}{\sigma^2}$, where μ_0^* and μ_1^* (and also j^*) are the values on which the maximum was obtained. Suppose that $f(x_k)$ is the actual value of y_k, and $\hat{f}(x_k)$ is the estimated value as in (13). Using (13) we get

$$|g_k - \hat{g}_k| \le S\sum_{i=j^*}^{k} \frac{6\kappa}{\sqrt{2n\delta_\kappa^*(x_i)}}.$$

Now we can formulate the following theorem.

Theorem 1. Let y_k be the set of observations of the signal $f(x)$, and $g_k, \hat{f}, \kappa, n, \delta_\kappa$ be defined as above. Then we have:

$$|g_k - \hat{g}_k| \le S\sum_{i=j^*}^{k} \frac{6\kappa}{\sqrt{2n\delta_\kappa^*(x_i)}}. \quad (14)$$

The main objective of this theorem is to show how an adaptive threshold in the presence of Gaussian noise can be obtained using constant threshold. For instance, in the case of slowdown detection described below, if we empirically selected some threshold on the smooth road, then the Theorem 1 provides an estimation of the adaptive threshold on "the real world road with arbitrary potholes and bumps".

6. TESTING

6.1 Description of the Test Track

We have designed a small autonomous vehicle model Fig.1, based on STM32F3-Discovery board containing the accelerometer and the gyroscope. The board was connected to a simple motor through the motor driver L293B. So we were able to control the direction of the vehicle movement. The power supply was provided by 4 1.2 V batteries or the USB connector.

We have implemented several software libraries in C: namely, the motor control, LED control, sensor, filter and fault detection algorithm libraries. Also we have implemented a simple program using this libraries to control the vehicle. Then we used OS Embox (high-modular operating system for resource-constrained devices) to program the board.

We set the accelerometer to 1344 HZ and read the values of the acceleration along x-axis in the polling mode. During the movement we stored obtained values in the flash memory. Then we restored this values from the flash memory on PC to see how the car behaves itself. Note that the vehicle is fully autonomous, and we use the stored data only to see the real values of accelerometer only for testing our method in Python. But we have the same C program inside the vehicle.

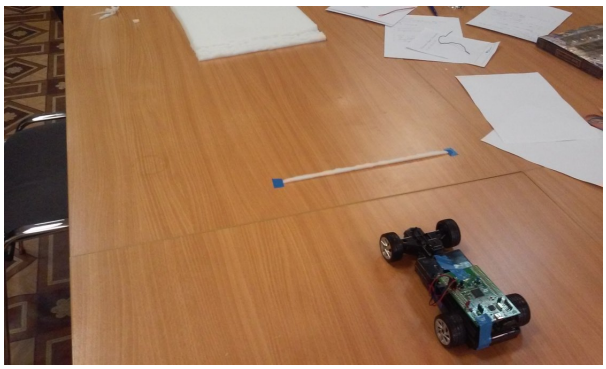

Fig. 1. Test track for the small car. It begins with a smooth road. The tiny piece of a foam rubber is a random disturbance. The big piece of a foam rubber is a serious obstacle.

When the car faces with a foam rubber it began to move slowly. And we want to recognize this change. Also you can see the tiny strip of a foam rubber. It is a random noise that is out of interest.

Our goal is to detect a foam rubber in real time, then stop the motor and turn the red LED on (Fig.2).

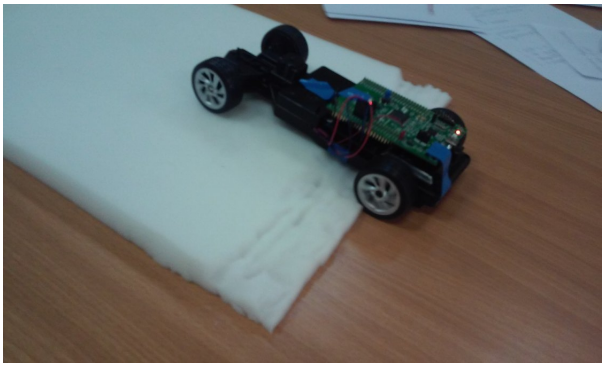

Fig. 2. The car detected a fault - motor has been stopped, red LED has been turned on.

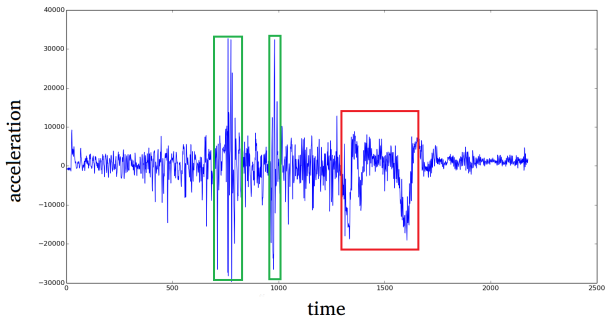

Fig. 3. Red rectangle denotes the moment when the fault occurred. Green rectangles denotes the moments when the car gone through the tiny piece of a foam rubber - it is out of interest.

6.2 The Experiment

On the Fig.3 we denoted with a red color the part that we want to recognize. One can see two green rectangles. The first one occurred when the front-wheel of the car gone through the tiny piece of a foam rubber. Similarly, another one was happened for the back-wheel of the car.

At first, we should to filter noise out of the acceleration values. Suppose we want to estimate the input signal at a given point x_0. We will use spatial adaptive estimation described above. Let $\kappa = 800$, and we calculate the ideal window Δ_k^* for every moment of time. To do this we use section "The idea" from Goldenshluger and Nemirovski (1997). Define the risk as follows

$$\rho_i = \kappa(n\delta_i)^{-1/2}.$$

And let us consider the following segments

$$D_i = [\hat{f}_i(x_0) - 2\rho_i, \hat{f}_i(x_0) + 2\rho_i].$$

Since the "dynamic" term $\omega_f(x_0, \delta_i)$ dominated by the "stochastic" term ρ_i, it follows that D_i, $i \leq i^*$ (i^* is an index of ideal window), have a point in common $f(x_0)$. So all that we need is to construct D_i iteratively while they intersect each other.

On the Fig.5 you can see how the window size (δ_i) changes during the time.

Now we want to detect a change in the mean value. But if you take a look at the graphic, you can see that the mean

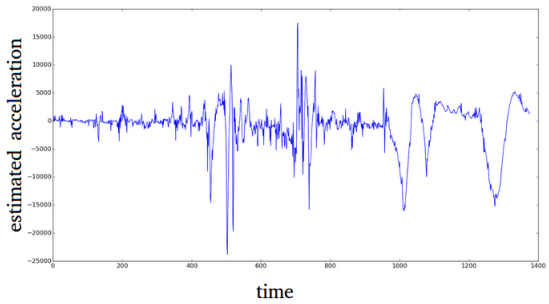

Fig. 4. Spatial adaptive estimation of the accelerometer data.

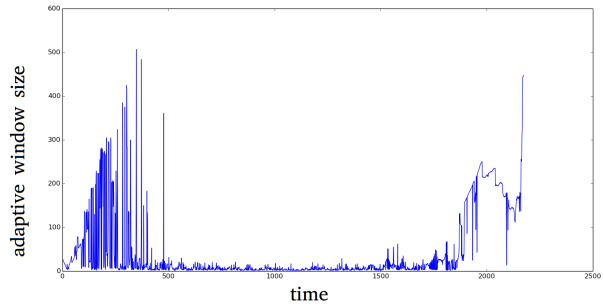

Fig. 5. Adaptive window size.

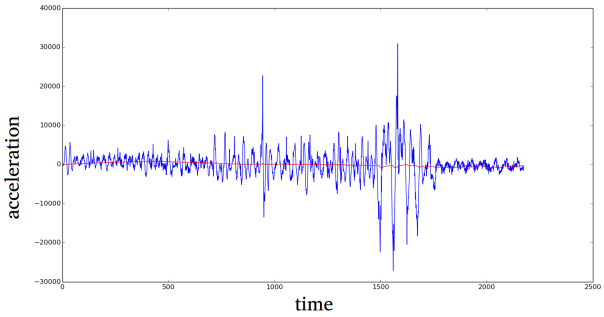

Fig. 6. Red line is the mean value of the accelerometer data (window size = 200).

value is about 0 anywhere. And it means that Generalized Likelihood Ratio is not appropriate in this case.

But if we take absolute values of the signal, then we get the following thing (Fig.7). One can see the abrupt change of the the mean value around $time = 1500$.

After we got the appropriate data (with a change in the mean value), we can apply Generalized Likelihood Ratio. We apply this to detect the change in mean value.

Then we select the threshold h for the signal without noise. Looking at the accelerometer values, we supposed that $h = 2000$. The calculated dynamic threshold is presented below. We need to note that we have use ϵ-set of the possible parameters of the θ. That is we have considered the set of the following type $\Theta := \{\theta_m = \theta_0 + m\epsilon : |\theta_m - \theta_0| \geq \nu\}$. So we have

$$g_k = \max_{1 \leq j \leq k} \sup_{\Theta} \sum_{i=j}^{k} \frac{y_i(\mu_0 - \mu_1)}{\sigma^2} + k\frac{\mu_1^2 - \mu_0^2}{2\sigma^2},$$

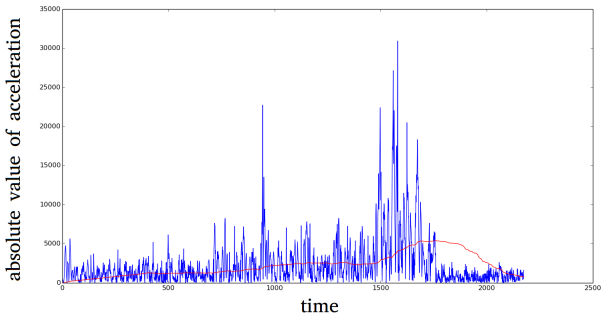

Fig. 7. Red line is a mean value of accelerometer data absolute values (window size = 200).

κ	percent of fails
400	60%
600	40%
700	30%
800	10%
900	10%
1000	10%
1500	20%
1600	30%
2000	60%

Table 1. Count of fails for different values of κ.

which can be easy calculated.

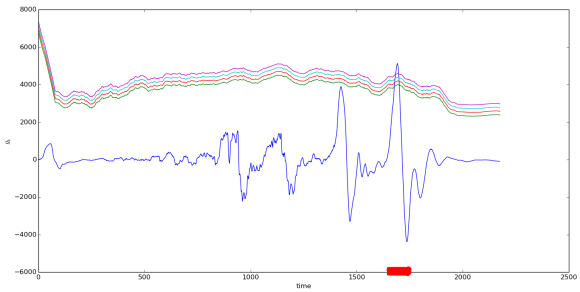

Fig. 8. Blue line is g_k, other lines denotes different dynamic thresholds. The region where a fault occurred is depicted by red color.

One can see a moment when the fault occurred - intersection of the green and blue lines. Also notice that if we would use constant $h = 2000$, then intersection would be occurred in the wrong place.

Also we provide the results of the experiments with different values of κ using 10 tracks of the accelerometer data for each value of κ. The best result was shown by $\kappa = 800,\ 900,\ 1000$. But it is difficult to calculate exactly which value of κ select in the concrete situation, because we don't know the signal-to-noise ratio. So it is selected empirically. In the Table 1 result are presented.

7. CONCLUSION AND FEATURE WORKS

The problem of abrupt change detection in autonomous systems is considered. The new method of the adaptive threshold estimation is offered. The experiment of the determining of the autonomous small car slowdown was made. The proposed dynamic threshold shows better capabilities than a constant threshold.

For the further investigations the estimation of the parameter κ in the conditions when the signal-to-noise ratio is unknown seems to be perspective. Also more complex description of a surface type could be useful. It could be done by hypothesis testing, where each hypothesis describes specific type of a surface.

7.1 Further Application

We want to consider the possible applications of this method for the medical diagnosis of biological processes in the human body. For instance, such diseases when the formation of a blood clot (thrombus) are accumulated inside a blood vessel.

It could be reasonable to accumulate some parameters of vessels like results of Ultrasound, and in case if there are many such results, one can to search them for a fault. Let us imagine that this results can be represented as a sequence of integers, so in this case we have to detect a changes in the mean value of this values. And if we will recognize a fault, it will help us to recognize a disease.

REFERENCES

Akimov, P. and Matasov, A. (2015). Weight and time recursions in dynamic state estimation problem with mixed-norm cost functions. *Automatic Control, IEEE Transactions on*, 60(4), 1050–1063.

Basseville, M. and Nikiforov, I.V. (1998). *Detection of Abrupt Changes: Theory and Application*. Prentice-Hall, Inc.

Blanke, M., Kinnaert, M., Lunze, J., and Staroswiecki, M. (2006). *Diagnosis and Fault-Tolerant Control*. Springer.

Davis, A., Nordholm, S., and Togneri, R. (2006). Statistical voice activity detection using low-variance spectrum estimatio and an adaptive threshold. *Audio, Speech, and Language Processing, IEEE Transactions on*, 14(2), 412–424.

Goldenshluger, A. and Nemirovski, A. (1997). On spatial adaptive estimation of nonparametric regression. *Mathematical methods of Statistics*, 6, 135–170.

Granichin, O., Zeev, V.V., and Toledano-Kitai, D. (2015). *Randomized Algorithms in Automatic Control and Data Mining*, volume 67. Springer.

Lepskii, O. (1990). On one problem of adaptive estimation in gaussian white noise. *SIAM Theory of Probability and Its Applications*, 35(3), 454–466.

Perhinschi, M., Napolitano, M., Campa, G., Seanor, B., Burken, J., and Larson, R. (2006). An adaptive threshold approach for the design of an actuator failure detection and identification scheme. *Control Systems, IEEE*, 14(3), 519–525.

Shu, L., Jiang, W., and Wu, Z. (2008). Adaptive cusum procedures with markovian mean estimation. *Computational Statistics & Data Analysis*, 52(9), 4395–4409.

Verdier, G., Hilgert, N., and Vila, J.P. (2008). Adaptive threshold computation for cusum-type procedures in change detection and isolation problems. *Computational Statistics & Data Analysis*, 52(9), 4161–4174.