SIGNAL SEGMENTATION FOR OPERATIONAL REGIMES DETECTION OF HEAVY DUTY MINING MOBILE MACHINES – A STATISTICAL APPROACH

Agnieszka WYŁOMAŃSKA

Hugo Steinhaus Center, Institute of Mathematics and Computer Science, Janiszewskiego 14a, Wrocław University of Technology, Poland
agnieszka.wylomanska@pwr.wroc.pl

Radosław ZIMROZ

KGHM CUPRUM Ltd CBR Sikorskiego 2-8, 53-659 Wrocław
Machinery Systems Division, Na Grobli 15, Wrocław University of Technology, Poland
rzimroz@cuprum.wroc.pl

Summary
In the paper an automatic procedure for diagnostic signals segmentation is proposed. The purpose of the procedures is to detect/identify part of the signal, that is related to stationary operation regime of machine. Detection and parameterization of such events might help to improve efficiency of machine usage, for example to minimize number of segments and their duration for machine operation under idle mode or overload mode. It is proposed to use a procedures that are based on statistical analysis to estimate the critical point of the division in a structural change in a time series. Two measures have been proposed for critical points detection: the first one is based on testing of empirical moment of order two for time subseries with length k, second one is related to analysis of second order moment moving along the signal. These techniques have been validated using simulations and then applied to real data acquired from on board monitoring system developed for mobile mining machines (loaders are considered here). Results of application are discussed in the paper.

Keywords: mobile machines, automatic operational regime detection, signal processing, signal segmentation

1. INTRODUCTION
Monitoring of machines and processes become a key approach to increase efficiency, quality and safety of production in companies in many branches of industry. Last decade provided many initiatives focused on implementation of such strategy to mining industry. It seems to be especially important in underground mining due to constraints related to environmental and safety issues. For example natural hazard, very non-comfortable working condition (temperature, humidity, noise etc)
motivates mining companies to use as much as possible automatic systems for monitoring and control of key elements of production chain in underground. Another issue related to increasing of production efficiency is machine performance monitoring and, in parallel, machine condition monitoring. To deal with this problems, a few years ago a kind of feasibility study on monitoring and localization of mobile mining machine in underground copper ore mine were done by [1]. Nowadays, many machines are equipped with On Board Diagnostic system, that is used to acquire and transmit via WiFi over 30 physical parameters. Collected data are transferred via underground IT infrastructure to Data Storage and Processing Center and after processing they are basis for daily reports about machine condition and performance.

Unfortunately, as it was mentioned, due to very harsh condition in underground mine, collecting signals should be followed by advanced techniques for signal validation, pre-processing and analysis (as reported by [2]). For both, performance and diagnostics context, parameters describing varying load/speed condition should be also monitored, understood and used for decision making [3-8]. Basically, monitoring of machine operation is used for:

- Identification of machine operation regimes
- Machine usage and effectiveness analysis
- Machine operator monitoring
- Condition monitoring and fault detection purposes

In order to fulfill these expectation, suitable physical variables should be acquired, transmitted, stored processed and visualized. In next section machine with proposed data acquisition system will be briefly described. In this paper, we will be focused on operational modes detection and signal segmentation according to detection results.

By operational mode it is understood here a periods of time when values of parameters describing operating condition are approximately constant. Such a problem has been noticed in other applications. It can be called trapping events (TE) detection. We will try to adapt and test TE detection method proposed in [9].

A main part of the paper will provide proposal of data processing procedure and results of its application to rotational speed signal. After segmentation basic statistical analysis will be done.

The rest of the paper is organized as follows. In section 2 we describe the mining machines. Next, in section 3 the procedures of finding trapping events and their lengths are presented. In section 4 we show the performance analysis of introduced methods for simulated data while the results for real signal are presented in section 5. Section 6 contains discussion and in section 7 we summarize the paper.

2. MINING MACHINES DESCRIPTION

In considered mining company there are more than 1000 heavy-duty mobile machines used for basic operation in mining production, namely drilling, bolting, loading and transporting processes (specialized trucks are used for short distance transport of materials from mining face to places when cyclic transport is replaced with belt conveyor system). Taking into account types and number of machines, loaders and haulage trucks seem to be the most important type of machines. In this paper data acquired during loader operations.

2.1. Machine description

Loader can be “decomposed” into several subsystems crucial from monitoring point of view. Two of them, namely drive system (engine and transmission system) and hydraulic system for lifting and steering the bucket seem to be the most important. To get information about condition and operation of these subsystems one might partially take advantage of build-in (provided by engine manufacturer) data acquisition system, unfortunately, other components should be equipped in sensors and data acquisition tools. So, concept deployed here is to use as much as possible existing data streams, add extra sensors and DAQ channels and finally integrate all into one output data stream.

Fig1. General view on loader used in underground mine

2.2. DAQ system description

Data acquired from the monitoring system cover temperatures, pressures, speeds (rotation of shaft and machine speed), torques, angles etc. Number and types of acquired variables depends on machine type. For loader analyzed here, the total number of output variables exceeds 30. Among others, a key variables used for analysis are machine speed, engine torque, temperatures and pressures of oil in engine, gearbox, hydraulic system. As it was said above, the monitoring system consists of two parts: devices acquiring set of variables related to engine operation (they are available via CAN data transmission protocol) and auxiliary monitoring system developed especially for monitoring purpose that allows to acquire rest of required variables.

In general, one might group all acquired data into 4 main class of signals: i) variables related to engine operation (torque, rotational speed, instantaneous fuel consumption), ii) temperatures, iii) pressures and iv) others [13]. Some auxiliary variables, as engine start/stop, total distance at given period, etc
are secondary variables, calculated based on primary ones mentioned above. In the next section, some examples of acquired signals will be presented and discussed.

3. PROCEDURE OF FINDING TRAPPING EVENTS AND THEIR LENGTHS

To the analysis of signal with trapping events (TE, called also traps) we propose to use two statistics. Both of them are based on the empirical second moment of underlying sample. But the first one, statistic \( R \), takes under consideration the empirical second moments of subsamples of examined process. The subsamples are constructed as consecutive vectors of observations with given length \( k \). More precisely, for given vector of observations \( X_1, X_2, \ldots, X_n \) and given length \( k \) the statistic \( R \) is defined as follows:

\[
R_j = \sum_{i=j}^{j+k} X_i^2 \quad j = 0, 1, \ldots, n - k
\]  

(1)

This statistic was proposed in [9] as an useful tool for recognition untypical behavior of real signal for which we observe the critical point after that statistical properties expressed in the language of second moment change. On the basis of the statistic it was also constructed the statistical test which helps in the problem of testing stationarity of given data.

In the problem of finding the moments of trapping events we use here the simple property of \( R \) statistic. If \( k < l \), where \( l \) is the critical change point, then we observe very characteristic behavior of \( R \) statistic, namely its mean value is constant for \( j \) smaller or equal than \( l - k \) and for \( j \) greater than \( l - 1 \) while otherwise it is a linear function with respect to \( j \). This property of \( R \) is discussed in [9]. On the basis of this property we can easily observe the point where the statistical properties of underlying signal change. Moreover we can also calculate the critical point. The statistic is useful for signals with visible change point but also for data where the point is difficult to find but we know that it exists. This kind of data were considered in [9] where authors examined the laboratory plasma fluctuations.

The second function, statistic \( C \), which is useful in the problem of finding the trapping events is a cumulative second moment of given sample. It means that for sample \( X_1, X_2, \ldots, X_n \) the \( C \) statistic is defined as follows:

\[
C_j = \sum_{i=1}^{j} X_i^2 \quad j = 1, 2, \ldots, n
\]  

(2)

Similar as the \( R \) statistic, the \( C \) function was considered in the problem of finding the critical point for data that can be divided into subsamples of different statistical properties, see [9]. The procedure of finding the critical point was based on the properties of \( C \) statistic. Namely for \( j \) smaller or equal \( 1 \) (1-the critical change point) the \( C \) statistic is a linear function with respect to \( j \) with zero slope while for \( j > 1 \), it is also a linear function with non-zero slope if in the signal the statistical properties (expressed in the language of second moment) change.

The statistics based on the cumulative moments (especially with even orders) are commonly used in practice. The empirical cumulative fourth moment was for example used in [10], where authors tested differences between distributions of two given samples, see also [11].

The method of finding the moment of trapping events and lengths of trapping events proceeds as follows:

1. For given sample calculate its increments and then - the \( R \) and \( C \) statistics according to formulas (1) and (2).
2. Calculate the differenced value of \( C \) statistic. Let us mention the differenced vector of given sample \( Z_1, Z_2, \ldots, Z_n \) is defined as follows:

\[
\Delta Z_j = Z_{j+1} - Z_j, \quad j = 1, 2, \ldots, n - 1
\]

3. Construct the zero-one vectors, \( RR \) and \( CC \), which indicate the moments where our statistics (\( R \) and differenced \( C \)) are equal zero (or approximately zero with given detection threshold \( e \)). The value equal to 1 means the vector of appropriate statistic is close (or equal) to zero.
4. On the basis of \( RR \) and \( CC \) statistics we define the starting points of moments of trapping events and lengths of trapping events. We assume the trap has length \( m \) if \( m \) consecutive observations are on the same level (or almost on the same level).

4. PERFORMANCE ANALYSIS BY SIMULATIONS

In this section we consider the simulated signal with visible trapping events. In Fig. 2 we present the simulated time series while in Fig. 3 – the differenced simulated signal.
The simulated signal is just a vector of uniformly (on the interval \([0,1]\)) distributed random variables in which we established the trapping events. For both cases, raw and differenced signals, locations of selected trapping events are marked by ellipses (Fig 2 and 3). The moments of trapping events and corresponding lengths as well as their levels are given in Table 1. Let us point out the level of each TE is different (constant value).

Table 1. Moments and corresponding lengths and levels of trapping events for simulated signal.

<table>
<thead>
<tr>
<th>Moment of trapping events</th>
<th>Length of trapping events</th>
<th>Level of trapping events</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>2</td>
<td>(1/7)</td>
</tr>
<tr>
<td>13</td>
<td>2</td>
<td>(2/7)</td>
</tr>
<tr>
<td>50</td>
<td>5</td>
<td>(3/7)</td>
</tr>
<tr>
<td>100</td>
<td>10</td>
<td>(4/7)</td>
</tr>
<tr>
<td>111</td>
<td>2</td>
<td>(5/7)</td>
</tr>
<tr>
<td>195</td>
<td>3</td>
<td>(6/7)</td>
</tr>
<tr>
<td>200</td>
<td>30</td>
<td>1</td>
</tr>
</tbody>
</table>

We consider the procedure for finding TE for differenced raw simulated signal. As it was mentioned, the procedure based on R statistic uses just the values of the statistic for differenced signal while the method based on the C statistic takes under consideration the differenced values of C statistic applied to differenced signal.

In Fig. 4 we present the R and differenced R statistic for \(k=2\) (estimated for increments of the signal). We observe the differentiation of R statistic does not change the values around zero then here we do not need to analyze the differenced values of this statistic.

As we observe, the large trapping events are clearly visible in both Fig. 3 and 4. Unfortunately the smaller ones are not seen (not detected) in the R and differenced R statistic. This comes directly from the definition of R. The length of detected trapping event is always shorter than the real one. The difference between length of TE calculated on the basis of R statistic and real length of TE is equal to parameter \(k\) that in our case is equal to 2. Moreover because we applied the R statistic to differenced signal then in Fig. 4 we actually observe the trapping events greater than 3. Because of this inconvenience in construction of RR vector that determines the moments of trapping events and their lengths, we increase lengths of TE adding two samples. In Fig. 5 we present the RR vector (after taking this correction). The larger TE are properly recognized unfortunately since TE shorter than 3 samples are not detected, they are not visible in Fig 5.

Next, we check our procedure of finding TE by using C statistic applied to differenced simulated signal presented. In Fig. 6 we show the C and differenced C statistic. Note that clearly visible in both statistics is the longest TE for \(t=200:230\).
In Fig. 7 we show the CC vector which is a result of the procedure of finding trapping events by using C statistic applied to differenced data. We remind the procedure actually is based on the differenced values of C statistic therefore in the procedure of finding trapping events we should take under consideration the correction, namely to the length of trapping events we have to add 1 sample. Moreover the moments of trapping events calculated on the basis of C statistic are shifted which should be also included in the analysis. In Fig. 7 we present the result after this correction. Similar, as for the procedure based on R statistic we do not observe here the trapping events of the length 2. After differencing the data we lose the small trapping events.

The results of the described procedures are presented in Table 2. This example shows the results based on R statistics are more reliable when this method is applied to the raw signal with larger trapping events (greater than 3) while the C statistic can be applied to data with trapping events greater than 2. Both methods work properly for data with different levels of trapping events.

<table>
<thead>
<tr>
<th>Moment of trapping events</th>
<th>Length of trapping events</th>
<th>R stat.</th>
<th>C stat.</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>50</td>
<td>5</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>100</td>
<td>10</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>111</td>
<td>2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>195</td>
<td>3</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>200</td>
<td>30</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

5. APPLICATION TO REAL DATA

5.1. Preliminary analysis

In this section application of proposed techniques to real data (namely engine speed) will be discussed. Engine Rotational Speed (shortly Eng_RPM) has been selected due to strong physical link between obtained results and knowledge about engine operation (A preliminary knowledge about idle mode speed and max. allowed speed is available). Fig 8 shows raw data and differenced raw data captured during several hours of machine operation. Again locations of TEs are marked by ellipses.

The purpose of application of proposed method is an automatic detection of trapping events i.e. parts of the signal when variation of amplitudes inside this segment is lower than assumed value (ideally zero). It can be interpreted as idle mode (if speed is around 700rpm) or continuous overloading (if speed is higher than max. allowed speed). Visual inspection allows to notice just a few such segments in raw data presented in Fig 8. In Fig. 9 we show the R (for k=2) and differenced R statistics applied for differenced raw signal while in Fig. 10 we present the C statistic and differenced C statistic applied to differenced raw signal. We observe in Fig. 9 (bottom/top) and 10 (bottom) visible TE for which the mentioned statistics have values around zero.
Fig. 9. R and differenced R statistic estimated for differenced raw real signal from Fig. 8 (note that R was estimated with $k=2$).

Fig. 10. C and differenced C statistic estimated for differenced raw real signal from Fig. 9.

Because for the real data (in contrary to simulated signal) the trapping observations are not exactly on the same level then in this case we should apply detection threshold. By observation of R statistic and differenced values of C statistic (for increments of raw signal) we decide to take the same detection threshold for both methods, namely $T=50$.

In Fig. 11 we present the value of R statistic (applied for differenced raw signal) with marked values smaller than the detection threshold while in Fig. 12 we show the differenced values of C statistic (applied for increments of raw signal) with marked values smaller than the given detection threshold.

It can be easily noticed that there are many samples with amplitudes smaller than threshold value, however, they cannot be classify as TEs. Our algorithm check also if adjacent samples are smaller than threshold.

Results of trapping events procedure detection using statistics R for $k=2$ and C are presented in Fig. 13, where we show RR and CC vectors (after mentioned appropriate corrections) that indicate at moments of trapping events and their lengths.

Fig. 11. R statistic estimated for differenced raw real signal with marked values smaller than detection threshold.

Fig. 12. Differenced C statistic estimated for differenced raw real signal with marked values smaller than detection threshold.

Fig. 13. RR and CC vectors for raw signal presented in Fig. 8.
Results presented in Fig 13 are not so clear (many detected TE of small values especially for C statistic) so further statistical analysis of detection results is recommended.

4.2. Analysis of detected trapping events using R statistic

Thanks to proposed procedure one can easily understand how many of trapping events have been occurred in the data. Moreover we can also conclude when trapping events have started/stopped, how long they are and finally on what level of engine rotational speed (see Fig 14).

From Fig 14 one might conclude that the most of TE are related to idle mode (~700rpm) and high speed mode (around 2000rpm). The next step is to aggregate information to more friendly form – to get this we perform further statistical analysis providing different useful information for maintenance staff (see Table 3).

![Fig.14. Trapping events detection using R statistic](image)

In Table 3 there are presented basic statistics regarding detected trapping events such as number of traps, median of moments of trapping events, mean of lengths of trapping events, standard deviation of lengths of trapping events, the smallest trap, the biggest trap, length of all trapping events and percentage of trapping data (for trapping events calculated by using R statistics).

<table>
<thead>
<tr>
<th>Parameters</th>
<th>R stat. for k=2</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of traps</td>
<td>88</td>
</tr>
<tr>
<td>Median of moments of TE</td>
<td>4001.5</td>
</tr>
<tr>
<td>Mean of lengths of TE</td>
<td>17.52</td>
</tr>
<tr>
<td>StD of lengths of TE</td>
<td>39.97</td>
</tr>
<tr>
<td>The smallest trap</td>
<td>4</td>
</tr>
<tr>
<td>The biggest TE</td>
<td>243</td>
</tr>
<tr>
<td>Length of all TE</td>
<td>1542</td>
</tr>
<tr>
<td>% of TE in signal</td>
<td>17%</td>
</tr>
</tbody>
</table>

4.3. Analysis of detected trapping events using C statistic

By analogy to previous section, the same data we analyze using C statistic. Fig 15 shows results for all trapping events while Fig 16 - for trapping events longer than 3 samples (to make possible comparison to R statistic which doesn’t detect short TEs). As for statistic R, the biggest number of trapping events is related to machine operation at idle mode and high speed mode.

![Fig.15. Trapping events detection using differenced C statistic](image)

It might be easily noticed that more trapping events are detected for C statistic than for statistic R. It happens for all TEs. When constraint (length of TE should be longer than 3) is applied, the number of detected for C statistic is still higher than for R statistic but comparable. Discussion regarding this observation will be provided in next section. Table 4 contains basic statistics regarding detected trapping events based on results obtained using statistic C.
Table 4. Basic data about detected trapping events using C statistic.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>C stat. (all TE)</th>
<th>C stat. (TE&gt;than 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of traps</td>
<td>311</td>
<td>95</td>
</tr>
<tr>
<td>Median of moments of TE</td>
<td>3755</td>
<td>3985</td>
</tr>
<tr>
<td>Mean of lengths of TE</td>
<td>7.16</td>
<td>16.62</td>
</tr>
<tr>
<td>StD of lengths of TE</td>
<td>22</td>
<td>38.5</td>
</tr>
<tr>
<td>The smallest TE</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>The biggest TE</td>
<td>243</td>
<td>243</td>
</tr>
<tr>
<td>Length of all TE</td>
<td>2538</td>
<td>1579</td>
</tr>
<tr>
<td>% of TE in signal</td>
<td>28%</td>
<td>17.9%</td>
</tr>
</tbody>
</table>

6. DISCUSSION

As we observe, both introduced statistics, namely C and R, are useful in the considered problem of finding moments of trapping events and their lengths. However as it was mentioned, the R statistic returns only traps greater than 3 while the method based on C statistic detects trapping events greater than 2. This is related to the construction of the statistics. It is worth to mention some important facts.

The parameter that can influence on the moments of trapping events (and their lengths) is the detection threshold (parameter e) mentioned in the description of the procedure of finding of moments of trapping events. This parameter influences on R as well as C statistic. Ideally it should be equal to zero but for real signal the trapping observations are not exactly on the same level. They are rather close to each other. In section 5 we chose arbitrary adequacy for R and C statistics which was equal to 50. This detection threshold was chosen by observing values of appropriate statistics. But in general the detection threshold should be selected by using more advanced techniques based on the probability property of examined data. This problem is an issue for further study.

The second parameter that influences the values of R statistic is the window length k. In our analysis we chose k=2 because the smaller k value the smaller trapping events can be detected.

In order to summarize the obtained results we have to indicate at differences between trapping events calculated by using R and C statistics. Here we compare the trapping events for R statistic for k=2 as this which in the better way describes the trapping events visible in the data. The first think is that the number of all trapping events for R statistic is smaller than for C statistic. It is related to the fact that in case of C statistic we observe large number of very small trapping events (equal 2, 3), for R statistic they are not observed. This is also visible in Fig. 13 where the large trapping events for method based on C statistic are interrupted by smaller trapping events. This phenomenon is not observed for method based on the R statistic. This is also confirmed by mean of lengths of trapping events for all traps for R and C statistics (see Table 3 and 4).

The second difference that should be pointed out is the lengths of all trapping events which is significantly different for R and C statistics. The last fact results also the difference in percentage of for all trapping observations. But if we consider only trapping events of lengths greater than 3 detected by using C statistic and trapping events obtained by R statistic we can confirm they coincide. In this case the number of traps is almost the same while the moments of trapping observations are very similar. Moreover lengths of trapping events also coincide. It results the percentage of trapping events in the signal is almost the same.

The final step could aggregate results from C and R to “amplify” correctness of TEs detection. It sounds reasonable when considering original data as random processes.

As a summary in Fig. 17 we show the moments of trapping events and their lengths recognized by using R and C statistic (for C statistic we took only trapping events greater than 3).

On the basis of Fig. 17 we can construct the simple procedure of finding trapping events on the basis of two statistics, R and C. Namely, we assume in the analyzed signal there is a trapping event at given moment if both methods indicate at this moment. Moreover the length of corresponding trapping event is a mean of trapping events detected for both methods. This simple procedure clearly indicates at trapping events detected by two examined statistics. This method allows for analyzing only trapping events that are detected together by two statistics. In Fig. 18 we show the result of the method which combines two statistics used to detection trapping events for real signal presented in Fig. 8.
REFERENCES


12. The I2Mine Website (Innovative Technologies and Concepts for the Intelligent Deep Mine of the Future), i2mine.eu


7. CONCLUSIONS

In the paper a problem of an automatic segmentation of time series has been discussed. Two criteria for segmentation have been proposed. They are based on second order moments.

First, we have tested our procedure using simulation with a priori known properties of signal, next we have applied segmentation procedure to real data, namely rotational speed of engine shaft in heavy duty mining machine.

The purpose of the paper is to identify parts of the signal with approximately constant values and perform statistical analysis of detected segments. Interpretation of the segment depends on mean value of samples inside segment. It could be classify as idle mode (machine operates without load) or overload mode. Theoretically it might happen that some segments will have mean speed belongs to wide range of speed.

Signal used for analysis covers several hours of machine operation. Results of TEs detection require further statistical analysis. It is expected to understand basic statistical measures - for example - as mean/median/min/max of TEs length.

It has been found that results obtained for two statistical criteria used for segmentation are slightly different. Finally, to make analysis more reliable, we decided to aggregate results for both statistics.

Finally it should be mentioned that this work has been developed in frame of I2Mine project (Innovative Technologies for the Intelligent Deep Mine of the Future [12]). Further work (threshold value estimation, using other variables etc) will be continued in the near future.
the assistant professor in the Institute of Mathematics and Computer Science at Wroclaw University of Technology. She is a member of the research group of Hugo Steinhaus Center for Stochastic Processes which goal is to organize, encourage and support research on and education in stochastic techniques as applied in science and technology. Her area of interest are anomalous diffusion processes, applications of stochastic processes in technical sciences and statistical analysis.

Radoslaw ZIMROZ, Ph.D. (2002), D.Sc. (2011), Professor at Wroclaw University of Technology (2012). His research activity is focused on signal processing and pattern recognition for condition monitoring of mining machines, especially working under time varying, non-stationary operating conditions. He is author more than 200 scientific works on such subject.