STUDIA INFORMATICA 2015
Volume 36 Number 3 (121)

Marcin MICHALAK
Silesian University of Technology, Institute of Informatics
Beata SIKORA
Silesian University of Technology, Institute of Mathematics
Jurand SOBCZYK
SOMAR S.A.

CORRELATION AND ASSOCIATION ANALYSIS IN WALL CONVEYOR ENGINES DIAGNOSIS

Summary. The paper presents the new way of machine diagnosis. The object of the research is the wall conveyor working in the coal mines. The work of the device was represented by three time series of current values of three conveyor's engines. Every startup of the conveyor work was described with almost twenty variables. The correlation analysis of over 3700 startups pointed interesting dependencies in the data. The association analysis gave sets of interpretable rules describing the proper way of conveyor work. The final prediction of the level of proper work is done on the basis of assumed number of last startups and their similarities to associations developed from the train data and represented by the association rules.

Keywords: machine diagnosis, correlation analysis, association analysis, association rules

ANALIZA KORELACJI I ASOCJACJI W DIAGNOSTYCE SILNIKÓW PRZENOŚNIKÓW ŚCIANOWYCH

Streszczenie. Artykuł przedstawia nowy sposób analizy pracy urządzeń – w tym przypadku przenośnika ścianowego w kopalni węgla kamiennego. Praca przenośnika opisana jest za pomocą trzech przebiegów poboru prądu przez każdy z silników. Każde uruchomienie przenośnika opisane zostało przez niemal 20 wskaźników, reprezentujących charakter i zmiany poboru prądu. Analizie poddano ponad 3700 uruchomień. Na podstawie analizy asocjacji wytypowano grupy reguł opisujących poprawny przebieg pracy przenośnika podczas uruchomienia. Końcowa ocena diagnostyczna polega na obserwacji w sposób ciągły historii uruchomień i porównaniu jakości opisu (dokładności reguł asocjacyjnych na obserwowanych przebiegach) z opisem uzyskanym z dostępnych wcześniej (historycznych, wzorcowych) danych.

Słowa kluczowe: diagnostyka urządzeń, analiza korelacji, analiza asocjacji, reguły asocjacyjne

1. Introduction

The correct work of so-called wall conveyors is very important for the proper operation of the longwall shearer. The conveyor task is to transport the coal mined by the harvester out of the longwall [1, 2]. For the longwall conveyor diagnostic it is important to analyze its startups. The conveyor is driven by three electric engines. One of them pulls and tensions a chain driving the conveyor, the other two just drive the conveyor chain [3, 4]. These engines make essential work and make it possible to transport the mined coal.

Due to the fact that the large amount of a coal is usually located on the wall conveyor, its startup is a critical moment. The conveyor drive system is usually equipped either with two-speed starting system (first gear - slow start of the conveyor; second gear - basic work of the conveyor) or with fluid coupling (by which startup and speed increase, that is the conveyor charging, occur gradually). This kind of coupling is offered inter alia by Voith, Transfluid, Siemens. The analysis of the startup and the correct work of the conveyor focuses on exceedances of assumed currents and the maximal currents during the work. The conveyor correct starting process consists of uniformly loading both main transporting engines. Uniform loading of the engines should manifest similar courses of time series reflecting the current drawn by the engines during their work.

The construction of the diagnostic models of machines/devices can be carried out as a planned experiment or can be based on the analysis of historical data. In the latter case we can have measurements which describe all states of the machine (including emergency states). Alternatively, we can have a certain subset of states (e.g. the state of proper operation of the machine). When we have only measurements which illustrate the state of proper operation of the machine, we can determine the model of this state and then observe whether the successive measurements are contained within this model. Going outside the model or observing a certain trend in changes can be a motive to raise the alarm. This type of diagnostics is applied for diagnosing the work of machines and devices which work in a normal production cycle of a plant where there is no time and no permission to run the planned experiments.

In the paper, the engines work was divided into two phases: startup and basic work. The startup phase is described by the maximum current value reached by the engines and the current rise time. The basic work phase is described by a time series illustrating the current drawn by the engines.

The purpose of this paper is to analyze the correlations and associations between parameters reflecting the process of the conveyor startup. The analysis used the Pearson linear correlation coefficient and the MagnumOpus program for induction of association rules [5, 6]. The analysis results can state a basis for the preparation of a diagnostic procedure verifying the correctness of the conveyor startups. The paper presents also the process of preparation, analysis and use of results flowing from it.

The paper is organized as follows: it starts from the presentation of the real dataset used in the research and the definition and interpretation of variables (indices) defined for the startup description. Then the results of the correlation analysis are presented with the division into groups of statistically significant correlations dealing with similar variables. Next part presents some association rules describing the correct diagnostic state of the conveyor work. Rules are also grouped by their premises to make their interpretation more easy. The goal of the paper is the presented in the following part the diagnostic procedure which may be helpful in the monitoring of the conveyor work. The paper ends with some final words and short description of the future works on the problem.

2. Background

The analysis of the conveyor work was performed on the data that described over eight weeks of the sampling through the DEMKop system [7]. Currents were sampled every second. In that time over 3700 single working times with startups were observed. As the startup the period of time between the first minimum in the current and the next maximum was defined.

From each working time the following indices were calculated:

- t – working time duration [s],
- m_1, m_2, m_3 – minimal value of the current during the startup [A],
- M_1, M_2, M_3 – maximal value of the current during the startup [A],
- $\Delta I_1, \Delta I_2, \Delta I_3$ – the increase of the current for each engine during the startup [A],
- v_1, v_2, v_3 – the speed of current increase for each engine [A/s],
- Δ – maximal difference between currents of engines in moments t_1, t_2 and t_3: $\max\{I_2(t_2) - I_1(t_1), I_2(t_2) - I_3(t_3)\}$ (the global difference) [A],
- δ – maximal difference between currents of engines during the startup (the local difference) [A],
- Δ_{12} – the maximal difference between currents of engines E_1 and E_2 in moments t_1 and t_2: $I_2(t_2) - I_1(t_1)$ (the global difference) [A],

- δ_{12} – the maximal difference between currents of engines E_1 and E_2 during the startup (the local difference) $[A]$,
- *est. lvl.* – "established level" at $t = 50$ $[s]$.

There are two "differences" defined, named as the local and the global. The global is the difference between currents remarking the end of the startup. For two engines it is just their difference. The local is defined in an analogous way but on samples coming from the same time for each engine. The visualization of the both of differences is shown in Figures 2 and 3. Additionally the interpretation of the current increase is also shown in Figure 1.

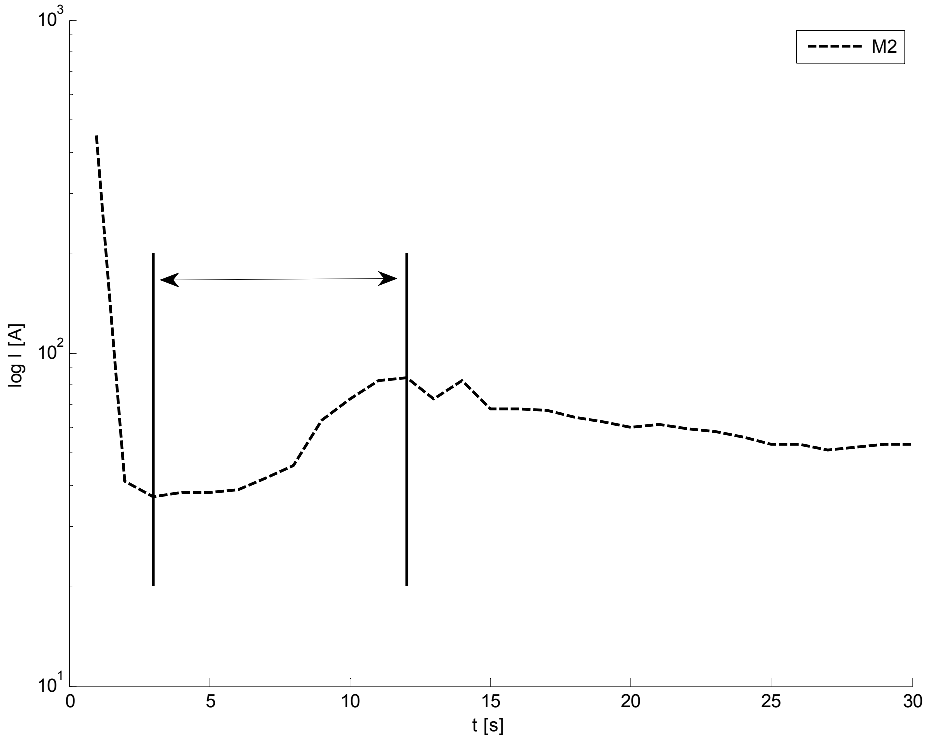

Fig. 1. The current increase time for the engine M_2
Rys. 1. Wzrost wartości poboru prądu przez silnik M_2

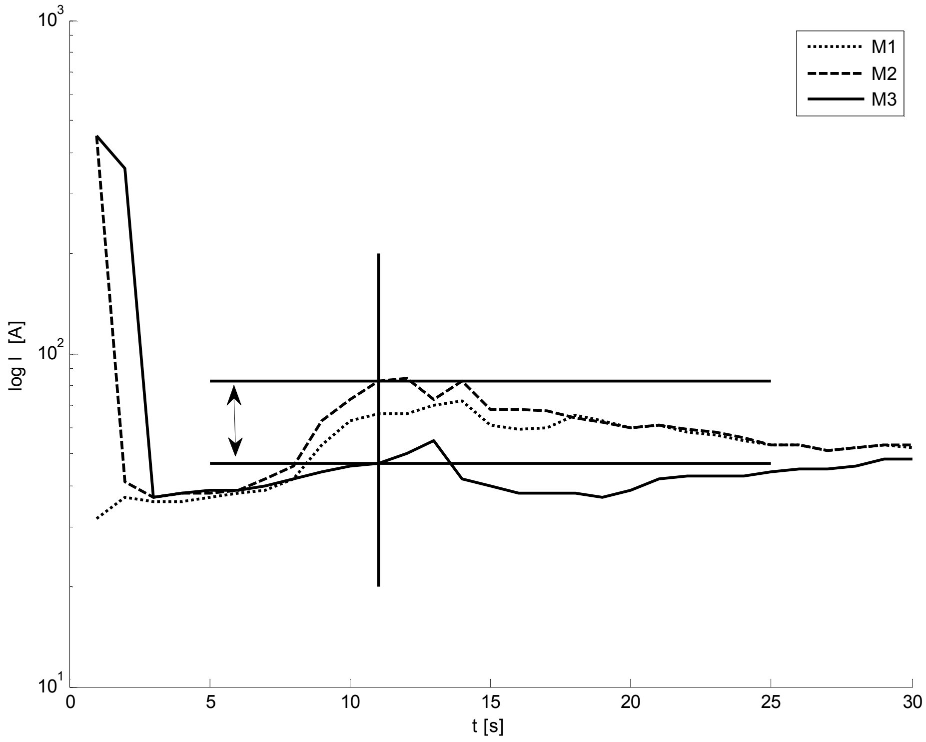

Fig. 2. The maximal local current difference between all engines
Rys. 2. Maksymalna lokalna różnica poboru prądu przez wszystkie silniki

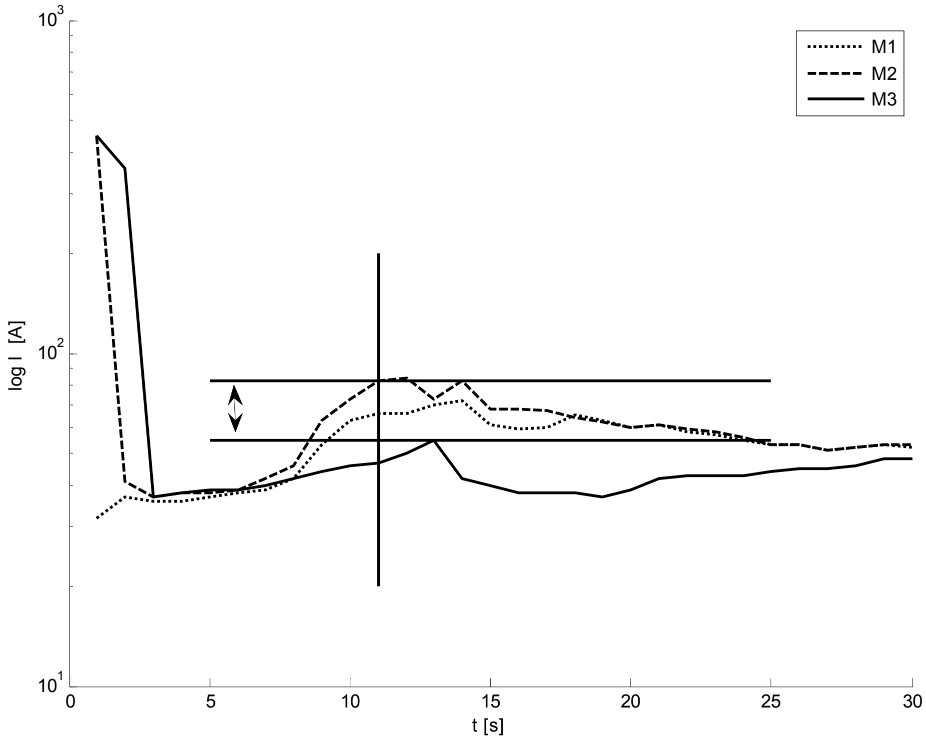

Fig. 3. The maximal global current difference between all engines
Rys. 3. Maksymalna globalna różnica poboru prądu przez wszystkie silniki

3. Correlation Analysis

The first step of the correlation analysis of features describing the conveyor startup consisted of calculation of Pearson correlation coefficients for each pair of startup features. The calculation results are presented in Table 1. Additionally, p-values of correlation coefficients were calculated and the results are presented in Table 2.

It is very common that high (or significant from the analyzed set point of view) values of the correlation coefficient do not have to represent sensible dependencies. Thus, it is recommended in the literature to take into consideration the correlation of those variables (features), which correlation can be at least partially explained. In this case the following correlations were found as interesting:

- correlation of minimal current of engines E_2 and E_3, equal to 0.63: $\rho(m_2, m_3) = 0.63$;
- correlations between the minimal current of engines E_1, E_2 and E_1, E_3 are rather small (smaller than 0.4) : $\rho(m_1, m_2) = 0.38$, $\rho(m_2, m_3) = 0.31$;
- maximal current of engines E_1 and E_2 (during the startup) are stronger correlated ($\rho(M_1, M_2) = 0.74$); those maximal currents (of E_1 and E_2) are weaker correlated with the engine E_3: $\rho(M_1, M_3) = 0.41$, $\rho(M_2, M_3) = 0.38$.

High values of the correlation coefficient were also observed between the current increase speed and the maximal current level during the startup:

- correlation coefficient of the E_1 engine current increase speed and the maximal current: $\rho(v_1, M_1) = 0.76$;
- correlation coefficient of the E_2 engine current increase speed and the maximal current: $\rho(v_2, M_2) = 0.88$;
- correlation coefficient of the E_3 engine current increase speed and the maximal current: $\rho(v_3, M_3) = 0.91$.

Also correlations between current increase and the maximal current of different engines become statistically significant. Values of coefficients are presented below:

- $\rho(v_2, M_1) = 0.65$,
- $\rho(v_1, M_2) = 0.74$.

Correlations of current increase times for all engines are rather small: from the range [0.33; 0.44].

Table 1

Pearson correlation coefficients

	t	m_1	M_1	ΔI_1	v_1	m_2	M_2	ΔI_2	v_2	m_3	M_3	ΔI_3	v_3	δ_{12}	Δ_{12}	δ	Δ	*est.lvl*
t	1.00	0.03	0.00	-0.01	0.00	0.03	0.08	0.01	0.06	-0.03	0.06	0.02	0.06	0.15	0.13	0.07	0.04	0.07
m_1	0.03	1.00	0.06	-0.64	0.20	0.38	0.07	-0.06	0.06	0.31	0.06	-0.05	0.05	-0.03	0.04	-0.01	0.02	0.10
M_1	0.00	0.06	1.00	0.28	0.76	0.00	0.74	0.37	0.65	-0.02	0.41	0.41	0.30	0.36	0.32	0.55	0.52	0.64
ΔI_1	-0.01	-0.64	0.28	1.00	-0.15	-0.22	0.24	0.44	0.11	-0.19	0.24	0.33	0.14	0.18	0.11	0.26	0.23	0.23
v_1	0.00	0.20	0.76	-0.15	1.00	-0.02	0.54	0.16	0.56	-0.04	0.27	0.24	0.22	0.26	0.23	0.37	0.34	0.46
m_2	0.03	0.38	0.00	-0.22	-0.02	1.00	0.08	-0.32	0.11	0.63	-0.02	-0.14	-0.05	0.03	0.08	-0.01	0.02	0.06
M_2	0.08	0.07	0.74	0.24	0.54	0.08	1.00	0.36	0.88	-0.04	0.38	0.41	0.26	0.53	0.60	0.61	0.64	0.60
ΔI_2	0.01	-0.06	0.37	0.44	0.16	-0.32	0.36	1.00	0.02	-0.23	0.36	0.42	0.23	0.20	0.22	0.36	0.36	0.35
v_2	0.06	0.06	0.65	0.11	0.56	0.11	0.88	0.02	1.00	-0.03	0.29	0.30	0.21	0.47	0.51	0.49	0.49	0.51
m_3	-0.03	0.31	-0.02	-0.19	-0.04	0.63	-0.04	-0.23	-0.03	1.00	0.07	-0.26	0.05	-0.05	-0.05	-0.04	-0.05	0.06
M_3	0.06	0.06	0.41	0.24	0.27	-0.02	0.38	0.36	0.29	0.07	1.00	0.30	0.91	0.15	0.14	0.46	0.44	0.55
ΔI_3	0.02	-0.05	0.41	0.33	0.24	-0.14	0.41	0.42	0.30	-0.26	0.30	1.00	-0.02	0.19	0.20	0.34	0.30	0.35
v_3	0.06	0.05	0.30	0.14	0.22	-0.05	0.26	0.23	0.21	0.05	0.91	-0.02	1.00	0.08	0.07	0.32	0.31	0.45
δ_{12}	0.15	-0.03	0.36	0.18	0.26	0.03	0.53	0.20	0.47	-0.05	0.15	0.19	0.08	1.00	0.83	0.63	0.55	0.22
Δ_{12}	0.13	0.04	0.32	0.11	0.23	0.08	0.60	0.22	0.51	-0.05	0.14	0.20	0.07	0.83	1.00	0.54	0.65	0.24
δ	0.07	-0.01	0.55	0.26	0.37	-0.01	0.61	0.36	0.49	-0.04	0.46	0.34	0.32	0.63	0.54	1.00	0.92	0.40
Δ	0.04	0.02	0.52	0.23	0.34	0.02	0.64	0.36	0.49	-0.05	0.44	0.30	0.31	0.55	0.65	0.92	1.00	0.40
est.lvl	0.07	0.10	0.64	0.23	0.46	0.06	0.60	0.35	0.51	0.06	0.55	0.35	0.45	0.22	0.24	0.40	0.40	1.00

Table 2

p-values of Pearson correlation coefficients

t	t	m_1	M_1	ΔI_1	v_1	m_2	M_2	ΔI_2	v_2	m_3	M_3	ΔI_3	v_3	δ_{12}	Δ_{12}	δ	Δ	*est.lv*
m_1	---	0.074	0.883	0.454	0.778	0.078	0.000	0.683	0.000	0.125	0.000	0.264	0.001	0.000	0.000	0.000	0.025	0.000
M_1	0.074	---	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.003	0.100	0.017	0.688	0.223	0.000
ΔI_1	0.883	0.001	---	0.000	0.000	0.926	0.000	0.000	0.000	0.314	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
v_1	0.454	0.000	0.000	---	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
m_2	0.778	0.000	0.000	0.000	---	0.331	0.000	0.000	0.000	0.028	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
M_2	0.078	0.000	0.926	0.000	0.331	---	0.000	0.000	0.000	0.000	0.152	0.000	0.001	0.047	0.000	0.484	0.353	0.000
ΔI_2	0.000	0.000	0.000	0.000	0.000	0.000	---	0.000	0.000	0.011	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
v_2	0.683	0.000	0.000	0.000	0.000	0.000	0.000	---	0.152	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
m_3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.152	---	0.080	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
M_3	0.125	0.000	0.314	0.000	0.028	0.000	0.011	0.000	0.08	---	0.000	0.000	0.004	0.004	0.002	0.008	0.005	0.000
ΔI_3	0.000	0.000	0.000	0.000	0.000	0.152	0.000	0.000	0.000	0.000	---	0.000	0.000	0.000	0.000	0.000	0.000	0.000
v_3	0.264	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	---	0.282	0.000	0.000	0.000	0.000	0.000
δ_{12}	0.001	0.003	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.004	0.000	0.282	---	0.000	0.000	0.000	0.000	0.000
Δ_{12}	0.000	0.100	0.000	0.000	0.000	0.047	0.000	0.000	0.000	0.004	0.000	0.000	0.000	---	0.000	0.000	0.000	0.000
δ	0.000	0.017	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.000	---	0.000	0.000	0.000
Δ	0.000	0.688	0.000	0.000	0.000	0.484	0.000	0.000	0.000	0.008	0.000	0.000	0.000	0.000	0.000	---	0.000	0.000
est.lvl	0.025	0.223	0.000	0.000	0.000	0.353	0.000	0.000	0.000	0.005	0.000	0.000	0.000	0.000	0.000	0.000	---	0.000
t	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	---

The biggest influence on differences between maximal current values of all engines has the second engine – to be more precise: its maximal startup current value. For this engine the correlation coefficient is $\rho(E_2, \delta) = 0.61$. The smaller influence is for the engine E_1 ($\rho(E_1, \delta) = 0.55$) and the smallest for the engine E_3 ($\rho(E_3, \delta) = 0.46$).

As it was presented before two different ways of measuring the difference between the maximal current values of all engines were defined. It occurs that there is a very strong correlation between these features amounting 0.92.

There is also a certain correlation between the established current value (considered as the current value after fifty seconds) and the maximal current values during the startup:

- $\rho(M_1, est.lvl) = 0.64$,
- $\rho(M_2, est.lvl) = 0.60$,
- $\rho(M_3, est.lvl) = 0.55$.

The correlation analysis may become the background of the more advanced association analysis, which is the subject of the next part of the paper.

4. Association Analysis

For the association analysis only 1000 from over 3700 startups were considered as the train-test samples. It was implied from the limitations of the Magnum Opus software. The precision measure was used as the rule quality evaluation method. It is the fraction of objects which fulfills the both rules sides (premises and a conclusion) – supporting objects, marked in Tables as supp – and a total number of objects fulfilling only premises of the rule – matching objects, marked in Tables as supp.

First set of association rules (Table 3) consists of ones that join the speed of current increase with the maximal current for the engine E_3. For higher values of speed of the current increase speed the maximal value of the engine E_3 becomes higher. These rules are quite strong so it can said that the speed of the current increase for the engine E_3 determines the maximal value of the current during the startup.

Table 3

First set of association rules

rule	precision	match	supp
IF $v_3 > 3.70$ THEN $M_3 > 79$	0.809	329	266
IF $2.50 \leq v_3 \leq 3.70$ THEN $66 \leq M_3 \leq 79$	0.656	340	223
IF $v_3 \leq 2.50$ THEN $M_3 \leq 66$	0.755	331	250

The similar set of rules for the engine E_1 (Table 4) consists of rules that are weaker than their equivalents for the engine E_3.

Table 4

Second set of association rules

rule	precision	match	supp
IF $v_1 \geq 6.00$ THEN $M_1 > 94$	0.786	238	187
IF $4.00 \leq v_1 \leq 6.00$ THEN $76 \leq M_1 \leq 94$	0.632	250	158
IF $2.92 \leq v_1 \leq 4.00$ THEN $64 \leq M_1 \leq 76$	0.596	260	155
IF $v_1 \leq 2.92$ THEN $M_1 \leq 64$	0.722	252	182

The weakest one says that for the current increase speed from the range [2.92; 4.0] the maximal current value is from the range [64; 76] but only 60% of the startups fulfils this relationship.

The same rules describing the work of the engine E_2 (Table 5) give more precise information. It is worth to be mentioned that for the range [3.0; 4.25] of current increase speed none of statistically significant association was found.

Table 5

Third set of association rules

rule	precision	match	supp
IF $v_2 \geq 6.13$ THEN $M_2 > 95$	0.806	248	200
IF $4.25 \leq v_2 \leq 6.13$ THEN $77 \leq M_2 \leq 95$	0.628	250	157
IF $v_2 \leq 3.00$ THEN $M_2 \leq 65$	0.736	258	190

Now, let us consider associations between minimal and maximal current values of all engines. From the correlation analysis we know that there were small correlations between minimal values of all engines. Association analysis confirms this observation.

In Table 6 we see rules describing only low values dependencies. Rules are also not strong. We just observe that minimal values of E_1 engine current determines minimal values of E_2 engine current.

Table 6

Fourth set of association rules

rule	precision	match	supp
IF $m_1 \leq 35$ THEN $m_2 \leq 36$	0.767	339	260
IF $m_2 \leq 36$ THEN $m_1 \leq 35$	0.583	446	260

Next table contains association rules generated from the maximal values of the currents. It occurs that only associations between engines E_1 and E_2 are statistically significant. We can see also that they are quite symmetric because they join values over 94 [A] and under 65 [A] in the analogical way. No association rule were detected for the middle levels of maximal current values.

Table 7

Fifth set of association rules

rule	precision	match	supp
IF $M_2 > 95$ THEN $M_1 > 94$	0.740	242	179
IF $M_1 > 94$ THEN $M_2 > 95$	0.725	247	179
IF $M_1 \leq 64$ THEN $M_2 \leq 65$	0.688	253	174
IF $M_2 \leq 65$ THEN $M_1 \leq 64$	0.682	255	174

Going back to the correlation analysis we see that maximal current values of engines E_1 and E_2 were not strongly correlated (0.56). That causes that rules build from these features are also quite weak, fulfilled only by 65% of objects (startups).

Two other associations (Table 8) are significant from the statistical point of view and join the speed of the current increase of two engines: E_1 and E_2.

Table 8

Sixth set of association rules

rule	precision	match	supp
IF $v_1 > 6.00$ THEN $v_2 > 6.13$	0.660	238	157
IF $v_2 > 6.13$ THEN $v_1 > 6.00$	0.633	248	157

It is another situation when we obtain set of symmetric rules.

Now some association rules describing the influence of the minimal and maximal current, its increase speed and the time on the maximal and minimal value of the current will be presented (from the statistical point of view only E_1 and E_2 engines were taken into consideration). The set of rules is presented in Table 9.

The rules presented above bind speed of the current increase and maximal current of different engines E_1 and E_2. Generally, the small current increase speed points the small maximal value of the current in the startup.

Table 9

Seventh set of association rules

rule	precision	match	supp
IF $v_1 > 6.00 \land v_2 > 6.13$ THEN $M_1 > 94$	0.898	157	141
IF $v_1 > 6.00 \land v_2 > 6.13$ THEN $M_2 > 95$	0.892	157	140
IF $v_1 < 2.92 \land v_2 < 3.00$ THEN $M_2 < 65$	0.820	150	123
IF $M_1 > 94 \land v_2 > 6.13$ THEN $M_2 > 95$	0.915	165	151
IF $M_2 > 95 \land v_1 > 6.00$ THEN $M_1 > 94$	0.923	155	143
IF $M_1 < 64 \land v_2 < 3.00$ THEN $M_2 < 65$	0.928	138	128
IF $M_2 < 65 \land v_1 < 2.92$ THEN $M_1 < 64$	0.891	147	131

Three very strong rules that bind similar dependencies for the engine E_3 were also discovered (Table 10).

Table 10

Eighth set of association rules

rule	precision	match	supp
IF $10 \leq \Delta I_3 \leq 12 \wedge v_3 < 2.50$ THEN $M_3 < 66$	0.891	147	131
IF $10 \leq \Delta I_3 \leq 12 \wedge 2.50 \leq v_3 \leq 3.70$ THEN $66 \leq M_3 \leq 79$	0.885	183	162
IF $10 \leq \Delta I_3 \leq 12 \wedge v_3 > 3.70$ THEN $M_3 > 79$	0.933	163	152

The next group of rules (Table 11) gives some information about the influence of the current increase speed and the different of the load on the maximal current of engines.

Table 11

Ninth set of association rules

rule	precision	match	supp
IF $M_1 > 94$ THEN $\delta > 41$	0.623	247	154
IF $v_2 > 6.13 \wedge \delta > 41$ THEN $M_2 > 95$	0.893	140	125
IF $v_1 > 6.00 \wedge \delta > 41$ THEN $M_1 > 94$	0.939	131	123
IF $v_2 > 6.13 \wedge \Delta > 33$ THEN $M_2 > 95$	0.935	138	129
IF $M_2 > 95 \wedge \delta > 41$ THEN $M_1 > 94$	0.824	142	117
IF $M_2 \leq 65 \wedge \Delta_{12} \leq 3$ THEN $M_1 \leq 64$	0.905	126	114
IF $M_1 \leq 64 \wedge \Delta_{12} < 3$ THEN $M_2 \leq 65$	0.983	116	114

We can observe that big values of maximal currents and high current increase speeds coexist with big current differences during the startup. It evidences of the unequal load of engines during the startup. The similar situation can be observed for small maximal currents which coexist with small differences between maximal currents of E_1 and E_2.

The next group of rules (Table 12) describes the influence of the current increase speed and the engine loading on the maximal current values.

Table 12

Tenth set of association rules

rule	precision	match	supp
IF $v_3 > 3.70 \wedge est.lvl > 70$ THEN $M_3 > 79$	0.963	135	140
IF $v_1 > 6.00 \wedge est.lvl > 70$ THEN $M_1 > 94$	0.943	123	116
IF $v_2 > 3.70 \wedge est.lvl > 70$ THEN $M_2 > 95$	0.900	130	117

The last set of rules (Table 13) confirm the strong correlation between two different ways of measuring the difference between all engines loading.

5. Diagnostic Procedure

Unfortunately (or fortunately, from the monitored conveyor point of view), during those eight weeks of observation none improper states of the conveyor were observed. It causes some problems of defining the diagnostic procedure. But even the proper work of the conveyor was not described in the exact (accurate) way as association rules precision were less than one.

Table 13

Eleventh set of association rules

rule	precision	match	supp
IF $\Delta_{12} \leq 3$ THEN $\delta_{12} \leq 2$	0.774	292	226
IF $\delta_{12} > 18$ THEN $\Delta_{12} > 15$	0.843	230	194
IF $\Delta < 9$ THEN $\delta < 12$	0.705	251	177
IF $18 < \Delta < 33$ THEN $24 < \delta < 41$	0.632	242	153
IF $\Delta > 33$ THEN $\delta > 41$	0.884	250	221

It brings the suggestion that the observation of the level of the current work description quality should be considered as the index of proper work. How the current work description quality should be understood? Let us consider the set of association rules describing dependencies between the current increase speed and maximal current for the first engine (E_1). This rules are repeated in Table 14.

Let us also assume that the “current work” is the period of 100 startups. Due to the value of v_1 each of these observed startup is recognised by exact one of the mentioned rules and increases the “left” rule's value. If this startup satisfies also the conclusion site of the rule then it also increases the ``both" value. This means that for observed startups for each rule also the precision on observed data can be calculated.

Table 14

Input rules for the diagnostic procedure

rule id	rule	precision	match	supp
1	IF $v_1 \geq 6.00$ THEN $M_1 > 94$	0.786	238	187
2	IF $4.00 \leq v_1 \leq 6.00$ THEN $76 \leq M_1 \leq 94$	0.632	250	158
3	IF $2.92 \leq v_1 \leq 4.00$ THEN $64 \leq M_1 \leq 76$	0.596	260	155
4	IF $v_1 \leq 2.92$ THEN $M_1 \leq 64$	0.722	252	182

As long as rules describe the work correctly (or in other words – conveyor behaves correctly) current precisions should not be smaller than on the train data. This situation is presented in Table 15.

Table 15
"Green level": correct startups during last 100 of them

rule id	precision	match	supp	curr. precision
1	0.786	10	9	0.900
2	0.632	15	12	0.800
3	0.596	35	25	0.714
4	0.722	40	35	0.875

The situation presented above can be interpreted as the "green level" of the first engine work correctness, joining the current increase and its maximal level. If at least one current precision of all rules is lower than the learnt precision it can be considered as the warning level ("yellow"). It can be caused by some incorrectness in the engine cooling system: for example the same maximal current value is reached much faster than in the correct state. The sample situation is presented in the Table 16.

Table 16
"Yellow level": the current precision of the third rule is smaller than the estimated from the train data

rule id	precision	match	supp	curr. precision
1	0.786	10	8	0.800
2	0.632	15	10	0.666
3	**0.596**	35	20	**0.571**
4	0.722	40	35	0.875

In the situation that all current precisions are smaller than the reference values the "alarm level" should be reported (Table 17).

In case of "red level" the operator of the device should give us the information whether there really was some problems with the cooling system or the device works correctly and the association rules should be re-learnt due to for example new work conditions of the conveyor.

Analogous analysis can be performed for other sets of association rules.

Table 17
"Red level": the current precision of the third rule is smaller than the estimated from the train data

rule id	precision	match	supp	curr. precision
1	**0.786**	10	6	**0.600**
2	**0.632**	15	7	**0.466**
3	**0.596**	35	15	**0.428**
4	**0.722**	40	27	**0.675**

6. Conclusions

In the paper the application of the correlation analysis and the association analysis for the evaluation of the conveyor diagnostic state. On the basis of the over eight weeks of the proper conveyor work observation some dependencies in the data were found. These dependencies are described as high values of statistically significant Pearson correlation coefficients and association rules.

On the basis of the association rules the new way of estimation of the diagnostic state of the wall conveyor was proposed. The obtained diagnostic procedure bases on the assumption that only the proper conveyor work is described and the deviation from this behavior is considered as the in-proper diagnostic state.

Because the given observations of the proper work made it possible to induct quite good association rules describing this diagnostic state, our next goal is to observe the same conveyor working under the same conditions in the proper and in-proper way. Induction of association rules describing two mentioned diagnostic states should give more precise models of making decisions about the state of the device.

Acknowledgements

The work was financially by Polish National Centre for Research and Development (NCBiR) grant PBS2/B9/20/2013 in frame of Applied Research Programs.

BIBLIOGRAPHY

1. Bartelmus W.: Condition Monitoring of Open Cast Mining Machinery. Wroclaw University of Technology Press, Wrocław 2006.
2. Siciński K., Isakow Z., Oset K.: Monitoring of Longwall Coal-cutting Machines' Work to Improve the Efficiency of the wall Extraction, Mechanization and Automation in the Mining Industry, 2002, Vol. 9, No. 381, p. 113÷120 (in Polish).
3. Michalak M., Sikora M., Sobczyk J.: Analysis of the Longwall Conveyor Chain Based on a Harmonic Analysis, Eksploatacja i Niezawodność – Maintenance and Reliability, 2013, Vol. 15, No. 4, p. 332÷336.
4. Gąsior S.: Diagnosis of Longwall Chain Conveyor, Przegląd Górniczy, 2001, Vol. 57, No. 7-8, p. 33÷36.

5. Webb G.: Discovering Associations with Numeric Variables, Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining, 2001, p. 383÷388.
6. Webb G.: Discovering Significant Patterns, Machine Learning, 2007, Vol. 68, No. 1, p. 1÷33.
7. DEMKop System: http://www.somar.com.pl/katalog-wyrobow/demkop,18,2,71

Omówienie

Artykuł przedstawia zastosowanie technik analizy korelacji i asocjacji w określaniu stanu diagnostycznego silników napędzających przenośnik ścianowy. Szczegółowej analizie zostały poddane początkowe fazy pracy przenośnika, to jest od momentu jego uruchomienia do osiągnięcia w miarę ustabilizowanego poziomu poboru prądu przez każdy z trzech silników. Faza ta jest szczególnie ważna z punktu widzenia bezpiecznej eksploatacji urządzenia, gdyż na ogół na przenośniku znajduje się spora ilość węgla, co utrudnia rozruch przenośnika i może powodować przeciążenia w poborze prądu.

W ramach badań przeanalizowano przebiegi poboru prądu wszystkich trzech silników w ponad 3700 poszczególnych uruchomianiach przenośnika. Każde uruchomienie zostało następnie opisane za pomocą kilkunastu parametrów, określających między innymi tempo narastania poboru prądu przez silniki, maksymalne pobory prądu, różnice w poborze prądu pomiędzy silnikami.

Tak przygotowane dane zostały następnie poddane analizie korelacji i asocjacji. Analiza korelacji wskazała między innymi na silne (i statystycznie istotne), a przede wszystkim interpretowalne w większości związki pomiędzy zmiennymi (tempo narastania poboru prądu a maksymalny pobór prądu, maksymalny pobór prądu a pobór w stanie ustalonym).

Drugą metodą analizy danych była analiza asocjacji. Z powodu ograniczeń oprogramowania indukcja reguł asocjacyjnych odbyła się jedynie na podstawie 1000 zebranych opisów uruchomień. W wyniku analizy zostało wybranych kilka grup reguł asocjacyjnych, wiążących ze sobą między innymi:

- prędkość narastania poboru prądu każdego silnika E_3 i jego maksymalny pobór prądu;
- minimalne wartości poboru prądu przez silniki E_1 i E_2,
- maksymalne wartości poboru prądu przez silniki E_1 i E_2,
- prędkości narastania poboru prądu przez silniki E_1 i E_2.

Na podstawie wyznaczonych grup reguł, a także wyliczonych dla nich miar jakości zaproponowano procedurę diagnostyczną, wykrywającą odstępstwo zachowania silników podczas kolejnych uruchomień od modelu regułowego. Ocena wyznaczona w wyniku tej procedury, dla każdej grupy reguł asocjacyjnych może przyjąć jedną z trzech wartości: „stan poprawny”, „stan ostrzegawczy” oraz „stan alarmowy”.

Adresses

Marcin MICHALAK: Silesian University of Technology, Institute of Informatics,
ul. Akademicka 16, 44-100 Gliwice, Poland, marcin.michalak@polsl.pl.
Beata SIKORA: Silesian University of Technology, Institute of Mathematics,
ul. Kaszubska 23, 44-100 Gliwice, Poland, beata.sikora@polsl.pl.
Jurand SOBCZYK: SOMAR S.A., ul. Karoliny 4, 40-186 Katowice, Poland,
j.sobczyk@somar.com.pl