



# Wykorzystanie reguł asocjacyjnych do analizy pracy wybranego urządzenia w oddziale wydobywczym

Agnieszka TRZCIONKOWSKA<sup>1)</sup>, Edyta BRZYCHCZY<sup>2)</sup>

<sup>1)</sup> Mgr inż.; AGH University of Science and Technology, Kraków, Mickiewicza 30, 30-059, Poland;  
email: agnieszka.trzcionkowska@agh.edu.pl

<sup>2)</sup> Dr hab. inż.; AGH University of Science and Technology, Kraków, Mickiewicza 30, 30-059, Poland

DOI: 10.29227/IIM-2016-02-13

## Streszczenie

W artykule przedstawiono możliwości wykorzystania zaawansowanych technik eksploracji danych do analizy pracy aparatury zabezpieczeniowej stacji transformatorowej, używanej w oddziałach wydobywczym w kopalniach podziemnych. Celem analizy było znalezienie czynników, które sprzyjają występowaniu zdarzenia zabezpieczeń oraz określenie, które z nich występują razem w przypadku zadziałania zabezpieczenia. W badaniach wykorzystano reguły asocjacyjne. Obliczenia zostały przeprowadzone z wykorzystaniem środowiska R oraz dodatku Rattle (Graphical User Interface for Data Mining in R). W wyniku przeprowadzonych analiz uzyskano reguły wskazujące pojawiające się zależności warunkujące pracę aparatury zabezpieczeniowej wybranej stacji transformatorowej.

Słowa kluczowe: analiza danych, górnictwo, reguły asocjacyjne, język R, stacja transformatorowa

## Wprowadzenie

W dobie społeczeństwa informacyjnego i gospodarki opartej na wiedzy, istotnego znaczenia nabiera jakość informacji, na podstawie których są podejmowane decyzje w przedsiębiorstwie. Ogromna ilość danych, które przechowywane są w stale rozrastających się bazach danych przedsiębiorstw zawiera zarówno tzw. wiedzę jawną jak i wiedzę ukrytą. Ta ostatnia, często poprzez brak procesu jej odkrywania pozostaje niewykorzystana, a może ona znacząco wpływać na jakość podejmowanych decyzji, i w konsekwencji, na przyszły stan przedsiębiorstwa. W procesie odkrywania wiedzy z danych zastosowanie znajdują różne techniki i narzędzia, dzięki którym zarządzający mogą analizować duże zbiory danych, w zależności od sformułowanego zadania eksploracji (np. opis, grupowanie, klasyfikacja czy odkrywanie reguł).

Najbardziej newralgicznym elementem przedsiębiorstwa, w ujęciu analizy systemowej, jest proces produkcyjny, który jest środkiem do osiągnięcia założonych celów (produkcyjnych, finansowych itp.). Stąd istotną kwestią dla zarządzających jest analiza przebiegu tego procesu, a szczególnie diagnoza występujących nieprawidłowości i anomalii. Proces wydobywczy, jest specyficznym procesem produkcyjnym, zarówno z uwagi na warunki jego realizacji jak i charakter tego procesu (proces logistyczny), w którym istotną rolę odgrywa wyposażenie wykorzystywane do realizacji procesów pod ziemią, a szczególnie to pracujące w wyrobiskach ścianowych.

Jednym z elementów koniecznych do realizacji robót w wyrobisku ścianowym jest stacja transformatorowa, która zasila sieci elektroenergetyczne

w oddziałach wydobywczym. Jej poprawna praca jest niezbędna do właściwej realizacji zaplanowanych do wykonania robót.

W artykule zaprezentowano przykład analizy pracy aparatury zabezpieczeniowej stacji transformatorowej w przykładowym wyrobisku ścianowym z wykorzystaniem reguł asocjacyjnych. Celem badań było wykazanie zależności, pojawiających się w analizowanym okresie pracy tej stacji, które powinny być przedmiotem zainteresowania zarządzających procesem wydobywczym.

## Reguły asocjacyjne

Analiza asocjacji wykorzystywana jest w celu znajdowania ukrytych zależności w dużych zbiorach danych w postaci prostych reguł. Polega ona na badaniu współwystępowania wartości różnych zmiennych. Jest to jedna z najczęściej wykorzystywanych technik analizy danych (szczególnie, w przypadku zmiennych o charakterze jakościowym). Pierwotnym zastosowaniem asocjacji była analiza danych transakcyjnych pochodzących z supermarketów, stąd też znana jest również pod nazwą analizy koszykowej [5].

Reguły asocjacyjne mają postać prostych zdań warunkowych typu: „jeżeli A to B”, co zapisuje się w postaci:

$$A \Rightarrow B$$

Element A nazywany jest poprzednikiem, natomiast B następnikiem w danej regule.

Jakość reguł określana jest przez trzy współczynniki:

- współczynnik wsparcia (ang. support),
- poziom zaufania (ang. confidence),

- przyrost (ang. lift).

Współczynnikiem wsparcia określa się prawdopodobieństwo, że dany przypadek zawiera dany element. Jest to również łączne prawdopodobieństwo (względna częstość współwystępowania) dla dwóch lub trzech elementów. Stąd wsparcie reguły na zbiorze przykładów  $P$  jest określone jako stosunek liczby przykładów ze zbioru  $P$  pokrywanych jednocześnie przez elementy  $A$  i  $B$  do liczby wszystkich przykładów w tym zbiorze [2], czyli:

$$w_{R(A \Rightarrow B)} = \frac{|P_{A \wedge B}|}{|P|}$$

Poziom zaufania reguły oznacza warunkowe prawdopodobieństwo, że obserwacja zawierająca element  $A$ , zawiera również  $B$ . Współczynnik ten na zbiorze na zbiorze przykładów  $P$  określa się jako stosunek liczby przykładów ze zbioru  $P$  pokrywanych jednocześnie przez elementy  $A$  i  $B$  do liczby przykładów z tego zbioru pokrywanych przez element  $A$  [2], czyli:

$$z_{R(A \Rightarrow B)} = \frac{|P_{A \wedge B}|}{|P_A|}$$

Natomiast przyrost jest to miara określająca czy fakt wystąpienia jednej wartości powoduje zwiększenie się prawdopodobieństwa wystąpienia drugiej, co można zapisać jako [10]:

$$p_{R(A \Rightarrow B)} = \frac{|P_{A \wedge B}|}{|P_A| \cdot |P_B|}$$

Przyrost przyjmuje wartości dodatnie. Gdy przyrost przyjmuje wartość 1 oznacza to, iż wystąpienie dwóch wartości nie wpływa na siebie, jeśli jest mniejsze od 1, wartości wykluczają się, a gdy przyrost jest większy od 1 to dwie wartości dopełniają się.

Najbardziej pożądanymi regułami są te o najwyższym współczynniku wsparcia, które są jednocześnie regułami silnymi (z wysokim poziomem zaufania).

Istnieje kilka algorytmów określania reguł (np. A priori, Carma, GRI), z których najbardziej rozpowszechnionym zarówno w wersji pierwotnej jak i zmodyfikowanej jest algorytm A priori.

Proces odkrywania reguł asocjacyjnych składa się najczęściej z dwóch etapów[1]:

1. Wyznaczane są wszystkie częste zbiory zdarzeń (czyli zbiory zdarzeń, które występują przy najmniej pewną, przyjętą minimalną liczbę razy).

2. Na podstawie częstych zbiorów zdarzeń tworzone są reguły, które spełniają warunek minimalnego wsparcia i poziomu zaufania.

Wspomniany wcześniej algorytm A priori korzysta z właściwości A priori, aby zmniejszyć przestrzeń rozwiązań. Właściwość ta stwierdza, że jeżeli dany zbiór zdarzeń  $Z$  nie jest częsty, to dodanie dowolnego zdarzenia  $A$  do zbioru  $Z$  nie spowoduje, że zbiór ten będzie zbiorem częstym. Właściwość ta redukuje znacząco przestrzeń poszukiwań dla tego algorytmu [7].

Dużą zaletą analizy asocjacji jest również możliwość graficznego przedstawienia otrzymanych związków pomiędzy badanymi elementami w postaci tzw. wykresów sieciowych.

### 3. Charakterystyka analizowanej stacji transformatorowej

Głównym zadaniem stacji transformatorowej jest zasilanie sieci elektroenergetycznych w oddziałach wydobywczych. Odbывается to poprzez przekształcanie, tj. najczęściej obniżanie, wartości napięcia dostarczanego do oddziału. Głównym elementem wyposażenia elektrycznego stacji jest transformator oraz aparatura łączeniowa, zabezpieczeniowa i sterująca.

Stacja transformatorowa przeznaczona jest do zasilania sieci niskonapięciowych o napięciu znamionowym wtórnym 3300 V. Napięcie znamionowe pierwotne, strony górnej wynosi 6000 V. Stacja ta wyposażona jest w transformator suchy trójfazowy o mocy znamionowej 2600 kVA.

Schemat ideowy analizowanej stacji transformatorowej przedstawiono na rysunku 1.

Stacja transformatorowa składa się z następujących elementów [8]:

- Komory górnego napięcia (GN), wyposażonej w zabezpieczenie nadmiarowo-prądowe,
- Komory transformatora (T),
- Komory dolnego napięcia (DN), wyposażonej w zabezpieczenie upływowo, nadmiarowo-prądowe oraz przyłączowe.

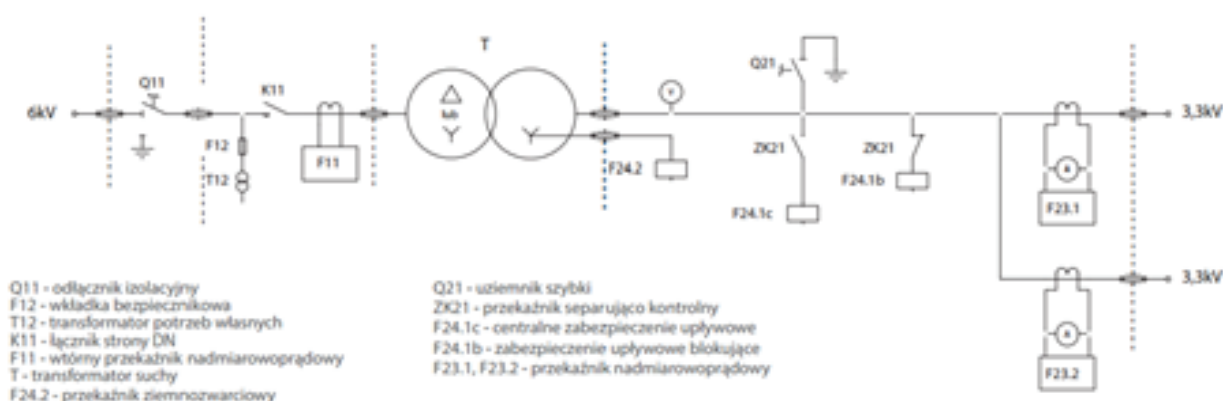
Po stronie komory DN stacja wyposażona jest w dwa odpływy. Każdy z odpływów wyposażony jest w zabezpieczenie nadmiarowo-prądowe, chroniące od skutków zwarcia, przeciążenia i asymetrii obciążeń. Ochrona od skutków pracy przy obniżonej lub uszkodzonej izolacji zasilanej instalacji realizowana jest przez zastosowanie zabezpieczenia upływowego- blokującego oraz centralnego.

Zabezpieczenia nadmiarowo-prądowe wyposażone są w trzy człony:

1. zwarciovyy – chroniący przed skutkami zwarc. Zadziałanie zabezpieczenia następuje przy wzroście wartości prądu w chronionym obwodzie powyżej prądu znamionowego pomnożonego

Rys. 1. Schemat ideowy stacji transformatorowej, źródło: [8]

Fig.1. Scheme of transformer station



przez krotność prądu nastawczego,

2. przeciążeniowy – chroniący od skutków przeciążeń z nastawialną charakterystyką zależną,

3. asymetry – chroniący od skutków pracy niepełnofazowej.

Zabezpieczenie upływowo-blokujące przeznaczone jest do wykrywania obniżonej wartości rezystancji izolacji. Działanie ochronne polega na pomiarze rezystancji izolacji w nadzorowanej instalacji. W momencie spadku rezystancji poniżej wymaganej wartości powoduje przerwanie obwodu sterowania, uniemożliwia włączenie napięcia na uszkodzony odpływ [8].

Analiza pracy zabezpieczeń stacji transformatorowej może dostarczyć cennych informacji na temat przebiegu oraz bezpieczeństwa prowadzonych prac. Wyniki przeprowadzonych analiz w tym zakresie zaprezentowano w dalszej części artykułu.

#### 4. Charakterystyka zbioru danych

Zbiór danych Dane\_analiza\_zabezpiezen przedstawia zapis rocznej pracy zabezpieczeń stacji transformatorowej. Przygotowanie danych wykorzystanych w analizie składało się z kilku etapów. Pierwszy z nich polegał na selekcji, czyli wyborze tylko tych zabezpieczeń, które wykazywały aktywność oraz wskazaniu dni, w których występowało zadziałanie tych zabezpieczeń.

Podstawowe zmienne, które są zapisywane w bazie to TagID, SampleDateTime oraz Sample\_Value. Na ich podstawie utworzono dodatkowe zmienne:

- na podstawie zmiennej TagID: zmienne Nazwa\_zaborazOzn\_podzesp,
- na podstawie zmiennej SampleDateTime: utworzono Dzień\_tyg, Godzina, Miesiąc, Zmiana,
- na podstawie zmiennych SampleDateTime oraz Sample\_Value: zmienną Czas\_trwania i

Czas\_trwania\_dysk.

Dokonano również czyszczenia danych, które polegało na usunięciu przypadków występowania braków danych. Braki te wynikały najczęściej z utraty transmisji w przesyłach danych.

W efekcie tych działań utworzono zbiór danych obejmujący 94 przypadki i 11 zmiennych. Charakterystykę zmiennych przedstawiono w Tabeli 1.

Analiza asocjacji wymaga stosowania wyłącznie zmiennych jakościowych. W tym celu zmienna ilościowa Czas\_trwania (Tab. 1.) została poddana dyskretyzacji. Nowo utworzona zmienna Czas\_trwania\_dyskr, posiada sześć wartości, tj.: „pon. 2s”, „2s– 0s”, „10s–60s”, „1 min–5min”, „pow. 5min”. Dla przypadków zadziałania zabezpieczenia można zauważyć, iż najliczniejsze są przedziały, w których zmienna Czas\_trwania przyjmuje wartości powyżej 10s oraz 1-5min (Rysunek 2).

W tabeli 2 przedstawiono 10 pierwszych przypadków ze zbioru danych.

Analiza opisanego zbioru danych została przeprowadzona zgodnie z przyjętym modelem eksploracji danych, który został szczegółowo przedstawiony w rozdziale 5.

5. Analiza pracy aparatury zabezpieczeniowej z wykorzystaniem reguł asocjacyjnych

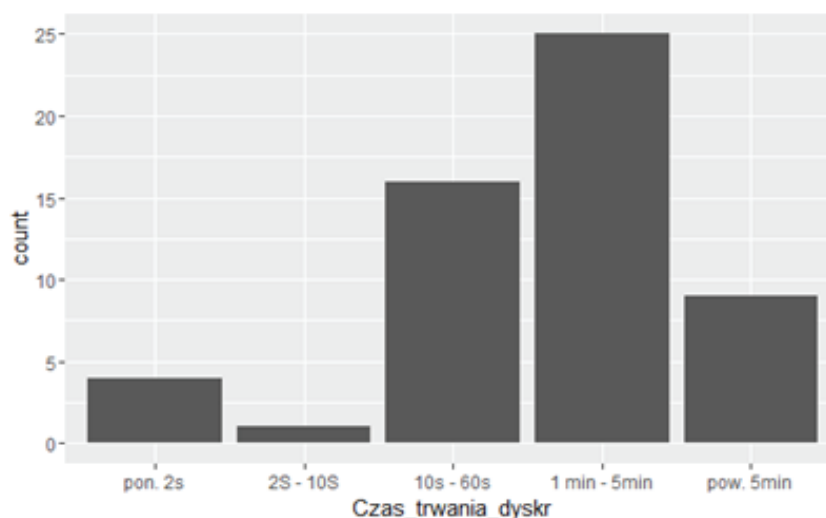
Do przeprowadzenia analizy przygotowanych danych przyjęto model eksploracji, przedstawiony na rysunku 3.

Do wstępnej analizy danych oraz wyznaczenia reguł asocjacyjnych wykorzystano środowisko obliczeniowe R. Jest to interpretowany język programowania oraz środowisko przystosowane do obliczeń statystycznych, jak również nowoczesnej wizualizacji wyników (w tym zakresie wykorzystano dodatek Rattle (Graphical User Interface for Data Mining in R)) [6, 9, 10].

Tab. 1. Charakterystyka zmiennych w zbiorze danych, źródło: opracowanie własne

Tab. 1. Characteristic of variables in a data set

Lp.	Oznaczenie zmiennej	Objaśnienie	Typ zmiennej	Możliwe wartości/jednostki
1	TagID	numer zabezpieczenia	jakościowa	12184,12186,12196,12204,12212,12214,12217
2	SampleDateTime	dokładny czas	czas	Przykładowo: „2014-11-24 15:30:02”
3	Sample_Value	zadziałanie zabezpieczenia (stan zabezpieczenia)	jakościowa	0- brak działania zabezpieczenia 1- załączenie zabezpieczenia
4	Nazwa_zab	rodzaj zabezpieczenia	jakościowa	Awaria Zabezpieczenie_zwarciove Zabezpieczenie_asymetrowe Urwanie_przetwornika
5	Ozn_podzesp	oznaczenie schematowe zabezpieczenia	jakościowa	zabezpieczenie górnego napięcia: <i>F11</i> zabezpieczenia na dwóch odpływach: <i>F23_1, F23_2</i>
6	Czas_trwania	czas trwania stanu zabezpieczenia	ilościowa	[s]
7	Dzien_tyg	nazwa dnia tygodnia	jakościowa	poniedziałek, wtorek, środa, czwartek, piątek, sobota, niedziela
8	Godzina	godzina zadziałania zabezpieczenia	jakościowa	0-24 godz.
9	Miesiac	miesiąc zadziałania zabezpieczenia	jakościowa	
10	Zmiana	oznaczenie pracującej zmiany	jakościowa	A,B,C,D
11	Czas_trwania_dyskr	zdykretyzowany czas trwania stanu zabezpieczenia	jakościowa	pon. 2s, 2s – 10s, 10s - 60s, 1 min - 5min, pow. 5min



Rys. 2. Histogram nowo utworzonej zmiennej Czas\_trwania\_dyskr dla przypadków zadziałania zabezpieczenia, źródło: opracowanie własne

Fig. 2. Histogram of newly created variable Czas\_trwania\_dyskr for protection activation cases

Wyniki kolejnych etapów eksploracji przedstawiono w kolejnych podrozdziałach.

### 5.1. Wstępna analiza danych

W celu zrozumienia struktury poszczególnych zmiennych oraz wykrycia relacji, które mogą ułatwić wybór zmiennych do dalszych analiz przeprowadzono tzw. eksploracyjną analizę danych.

W pierwszej kolejności sporządzono histogramy wybranych zmiennych jakościowych (Rys.4) Wykresy te sporządzono zarówno dla przypadków działania zabezpieczeń (Sample\_Value = 1, oznaczone kolorem niebieskim) oraz przypadków, w których nie działały zabezpieczenia (Sample\_Value = 0, oznaczone kolorem różowym). Największe różnice w histogramach można zauważyć

Tab. 2. Struktura bazy danych (10 przykładowych przypadków dla wybranych zmiennych), źródło: opracowanie własne

Tab. 2. Structure of data base (10 cases for selected variables)

Sample_Value	Nazwa_zab	Ozn_podzesp	Czas_trwania	Dzien_tyg	Godzina	Miesiac	Zmiana	Czas_trwania_dyskr
1	Awaria	F23_2	102.507	sobota	6	2015-01	A	1 min - 5min
0	Awaria	F23_2	2.093	sobota	6	2015-01	A	2s- 10s
1	Awaria	F23_2	71.540	sobota	6	2015-01	A	1 min - 5min
1	Awaria	F23_2	17.770	wtorek	22	2015-03	C	10s - 60s
1	Awaria	F23_2	25.010	środa	13	2015-03	B	10s - 60s
1	Awaria	F23_2	17.677	czwartek	2	2015-03	D	10s - 60s
1	Zabezpieczenie_asymetrowe	F23_2	331.360	czwartek	8	2014-12	A	pow. 5min
0	Zabezpieczenie_asymetrowe	F23_2	384.620	czwartek	8	2014-12	A	pow. 5min
1	Zabezpieczenie_asymetrowe	F23_2	80.684	czwartek	8	2014-12	A	1 min - 5min
0	Zabezpieczenie_asymetrowe	F23_2	349598.370	czwartek	8	2014-12	A	pow. 5min
1	Zabezpieczenie_zwarciove	F23_2	115.004	sobota	22	2015-01	C	1 min - 5min



Rys. 3. Przyjęty model eksploracji w analizie, źródło: opracowanie własne

Fig. 3. Model of data exploration

w przypadku zmiennych TagID, Czas\_trwania\_dyskr, Dzien\_tyg, Godzina, Zmiana. Różnice widać też dla zmiennej Nazwa\_zab.

Na podstawie przedstawionych wykresów można stwierdzić, iż wyróżnione zmienne są wskazane do dalszego wykorzystania, ponieważ występują różnice w ich rozkładach dla przypadków występowania zadziałania zabezpieczeń oraz braku takiego zadziałania (stan normalny).

### 5.2. Wybór zmiennych do analizy

Działania podejmowane na etapie wyboru zmiennych powinny przeciwdziałać powielaniu się informacji w wyznaczanym modelu (ma to istotny wpływ na jego jakość). Szczególnie ważne jest usunięcie zmiennych objaśniających, które wykazują istotne zależności pomiędzy sobą.

Po wstępnej analizie danych z wykorzystaniem histogramów, zbadano zależności występujące pomiędzy zmiennymi z wykorzystaniem tabel wielo-

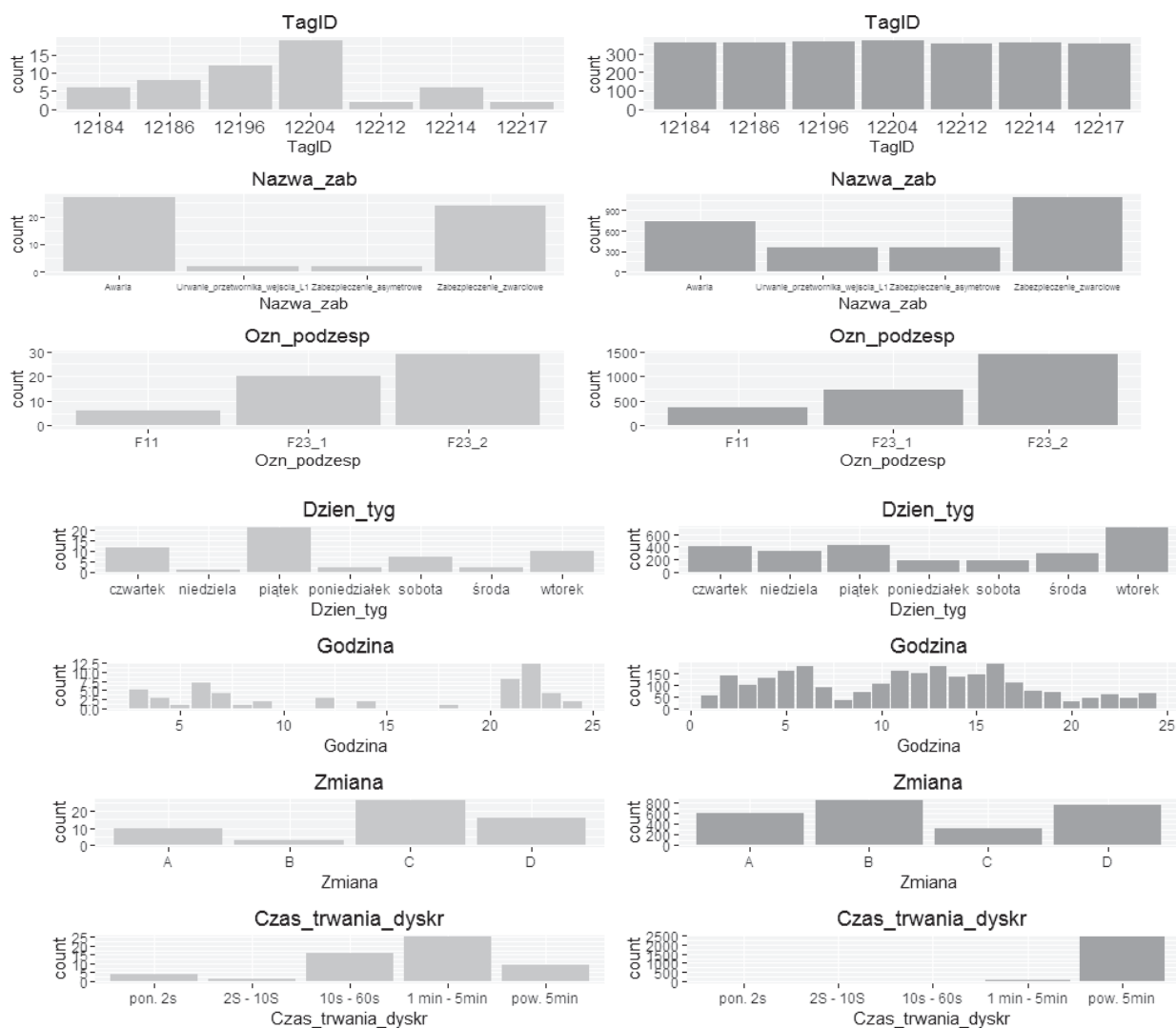
dzielczych. W wyniku przeprowadzonej analizy, jak należało oczekiwać, wskazano na zależności pomiędzy TagID i Nazwa\_zab (Rys.5) oraz Godzina i Zmiana (Rys.6).

Pierwsza zależność jest wynikiem utworzenia zmiennej Nazwa\_zab z TagID. Można zauważyć, iż poszczególne oznaczenia ID zabezpieczeń odpowiadają konkretnym nazwom zabezpieczeń. Druga zależność, przy ustalonej organizacji pracy w czasie doby jest raczej oczywista. Z punktu widzenia analizy pracy w oddziale wydobywczym, bardziej interesującym zagadnieniem może okazać się porównanie poszczególnych zmian niż konkretnych godzin występowania zadziałania zabezpieczeń. Stąd też w dalszej analizie pominięto zmienne Tag\_ID oraz Godzina.

### 5.3. Analiza z wykorzystaniem reguł asocjacyjnych.

Wyznaczenie reguł asocjacji przeprowadzono z wykorzystaniem biblioteki „arules”. Dostarcza ona infrastrukturę do analizowania, przetwarzania





Rys. 4. Histogramy zmiennych jakościowych. Po lewej – przypadki zadziałania zabezpieczeń, po prawej – przypadki, w których działa zabezpieczenie, źródło: opracowanie własne

Fig. 4. Histograms for qualitative variables (blue color – protection activation, rose color – time of protection operation)

i wizualizacji danych dotyczących transakcji oraz wzorców i reguł asocjacji [3].

Analizę rozpoczęto od wprowadzenia ustawień dla algorytmu tj.:

- wybór zbioru danych wykorzystanych do analizy (data),
- określenie minimalnego wsparcia oraz zaufania, a także minimalnej długości reguł (parameter),
- sprecyzowanie następników i poprzedników, a także wybór sposobu sortowania reguł (appearance).

Po otrzymaniu wstępnych, bardzo licznych reguł, dokonano wyboru tych najistotniejszych, cechujących się pożądanymi wartościami wsparcia i zaufania. Usunięto wszystkie reguły redundantne. Najważniejsze reguły sortowane były względem parametru przyrost (lift). Wyniki eksploracji reguł asocjacji dla wybranych ustawień algorytmu przedstawiono w kolejnych podrozdziałach.

### 5.3.1. Reguły dotyczące zmiennej “Sample\_Value=1”

W pierwszej kolejności szukano odpowiedzi na pytanie, jakie czynniki sprzyjały faktowi zadziałania zabezpieczenia (reguła z następnikiem: “Sample\_Value=1”). Jako parametry algorytmu wprowadzono minimalne wsparcie na poziomie 0,1, zaufanie na poziomie 0,4 oraz minimalną długość reguł jako 2. Wyznaczone reguły przedstawiono w tabeli 3.

Na podstawie uzyskanych wyników można stwierdzić, że czynnikami sprzyjającymi przypadkom zadziałania zabezpieczenia była zmiana C, dzień tygodnia piątek, miesiąc styczeń, zabezpieczenia Awaria oraz odpływ F23\_2. Współczynniki zaufania wyznaczonych reguł wynoszą ponad 50%.

Zaprezentowane reguły nie charakteryzują się zbyt wysokimi współczynnikami wsparcia czy zaufania, natomiast dostarczają bardzo cennych

Rys. 5. Tabela wielozdzielcza dla zmiennych TagID i Nazwa\_zab, źródło: opracowanie własne

Fig. 5. Cross Tab. for variables TagID and Nazwa\_zab

		TagID	12184	12186	12196	12204	12212	12214	12217
Nazwa_zab	Count								
	Total %								
Awaria	Count	0.00	371.00	0.00	393.00	0.00	0.00	0.00	0.00
	Total %	0.00	14.30	0.00	15.14	0.00	0.00	0.00	0.00
urwanie_przetwornika_wejscia_L1	Count	0.00	0.00	0.00	0.00	0.00	0.00	0.00	359.00
	Total %	0.00	0.00	0.00	0.00	0.00	0.00	0.00	13.83
zabezpieczenie_asymetrowe	Count	0.00	0.00	0.00	0.00	359.00	0.00	0.00	0.00
	Total %	0.00	0.00	0.00	0.00	13.83	0.00	0.00	0.00
zabezpieczenie_zwarciove	Count	367.00	0.00	379.00	0.00	0.00	367.00	0.00	0.00
	Total %	14.14	0.00	14.61	0.00	0.00	14.14	0.00	0.00

Rys. 6. Tabela wielozdzielcza dla zmiennych Godzina i Zmiana, źródło: opracowanie własne

Fig. 6. Cross Tab. for variables Godzina and Zmiana

		Zmiana	A	B	C	D
Godzina	Count					
	Total %					
1	Count	0.00	0.00	0.00	56.00	
	Total %	0.00	0.00	0.00	2.16	
2	Count	0.00	0.00	0.00	140.00	
	Total %	0.00	0.00	0.00	5.39	
3	Count	0.00	0.00	0.00	108.00	
	Total %	0.00	0.00	0.00	4.16	
4	Count	0.00	0.00	0.00	132.00	
	Total %	0.00	0.00	0.00	5.09	
5	Count	0.00	0.00	0.00	163.00	
	Total %	0.00	0.00	0.00	6.28	
6	Count	0.00	0.00	0.00	188.00	
	Total %	0.00	0.00	0.00	7.24	
7	Count	93.00	0.00	0.00	0.00	
	Total %	3.58	0.00	0.00	0.00	
8	Count	37.00	0.00	0.00	0.00	
	Total %	1.43	0.00	0.00	0.00	
9	Count	74.00	0.00	0.00	0.00	
	Total %	2.85	0.00	0.00	0.00	
10	Count	105.00	0.00	0.00	0.00	
	Total %	4.05	0.00	0.00	0.00	
11	Count	161.00	0.00	0.00	0.00	
	Total %	6.20	0.00	0.00	0.00	
12	Count	153.00	0.00	0.00	0.00	
	Total %	5.90	0.00	0.00	0.00	
13	Count	0.00	182.00	0.00	0.00	
	Total %	0.00	7.01	0.00	0.00	
14	Count	0.00	137.00	0.00	0.00	
	Total %	0.00	5.28	0.00	0.00	
15	Count	0.00	147.00	0.00	0.00	
	Total %	0.00	5.66	0.00	0.00	
16	Count	0.00	189.00	0.00	0.00	
	Total %	0.00	7.28	0.00	0.00	
17	Count	0.00	112.00	0.00	0.00	

informacji, umożliwiających charakteryzowanie parametrów sprzyjających wystąpieniu zadziałania zabezpieczenia i stanowiących podstawę do dalszych pogłębionych analiz.

### 5.3.2. Reguły dotyczące przypadków zadziałania zabezpieczeń

Kolejna analiza miała na celu odpowiedź na pytanie, jakie parametry charakteryzują działanie zabezpieczeń. Skupiono się w niej na badaniu współwystępowania wartości różnych zmiennych dla przypadków zadziałania zabezpieczenia. Jako parametry algorytmu wprowadzono minimalne wsparcie na poziomie 0,3, zaufanie na poziomie 0,7 oraz minimalną długość

reguł jako 2. Wyznaczone reguły przedstawiono w tabeli 4.

Na podstawie otrzymanych wyników można stwierdzić, że wszystkie awarie w styczniu na zmianie C pojawiły się w piątek. Ponadto 72% przypadków zadziałania zabezpieczenia o czasie trwania tego stanu od 1–5 min zdarzyło się w miesiącu styczniu oraz ponad 70% awarii wykazywanych było na odpływie F23\_2.

Na podstawie wskazanych zależności można sformułować dalsze pytania, na które odpowiedzi mogą być interesujące dla zarządzających procesem (np. co działo się w styczniu, jakie roboty wykonywane są najczęściej w piątki, jak wykonywane są prace przez zmianę C itp.). War-

Tab. 3. Reguły dotyczące zmiennej "SampleValue=1" jako następnik, źródło: opracowanie własne

Tab. 3. Rules for variable "SampleValue=1" as consequent

Reguły	Wsparcie	Zaufanie	Przyrost
{Nazwa_zab=Zabezpieczenie_zwarciove} => {SampleValue=1}	0.255	0.615	1.052
{Ozn_podzesp=F23_2} => {Sample_Value=1}	0.309	0.604	1.033
{Ozn_podzesp=F23_1} => {Sample_Value=1}	0.213	0.556	0.949
{Nazwa_zab=Awaria} => {Sample_Value=1}	0.287	0.551	0.942
{Zmiana=C} => {Sample_Value=1}	0.277	0.542	0.926
{Miesiac=2015-01} => {Sample_Value=1}	0.287	0.540	0.923
{Dzien_tyg=piątek} => {Sample_Value=1}	0.223	0.512	0.875

Tab. 4. Przykłady reguł dotyczących tylko przypadków zabezpieczenia, źródło: opracowanie własne

Tab. 4. Examples of rules for protection activation

Reguły	Wsparcie	Zaufanie	Przyrost
{Nazwa_zab=Awaria,Miesiac=2015-01,Zmiana=C} => {Dzien_tyg=piątek}	0.345	1.000	2.619
{Czas_trwania_dyskr=1 min - 5min} => {Miesiac=2015-01}	0.327	0.720	1.467
{Nazwa_zab=Awaria} => {Ozn_podzesp=F23_2}	0.345	0.704	1.335

to zwrócić również uwagę na fakt, iż zadziałanie zabezpieczenia to sytuacja skrajna, wskazująca na wystąpienie stanu niebezpieczeństwa. Wszelkie informacje umożliwiające, chociaż w niewielkim stopniu, przewidzenie w jakich warunkach może ono wystąpić są bardzo cenne.

### 5.3.3. Reguły dotyczące sprecyzowanego konkretnego poprzednika

Aby przeprowadzić bardziej szczegółową analizę dotyczącą interesujących wartości zmiennych, można wyznaczyć reguły ze wskazaniem poprzednika. I tak, aby przyjrzeć się, co dzieje się na wybranych zmianach - zmianie C (wskazanej we wcześniejszej analizie) oraz zmianie A (na której prowadzone są najbardziej intensywne roboty eksploatacyjne), wskazano te wartości jako poprzedniki w regułach. Jako parametry algorytmu wprowadzono minimalne wsparcie na poziomie 0,3, zaufanie na poziomie 0,7 oraz minimalną długość reguł jako 2. Wyznaczone reguły przedstawiono w tabeli 5.

Na ich podstawie można wnioskować, że ponad 76% przypadków zadziałania zabezpieczeń na

zmianie C było w piątek, w ponad 80% w styczniu i głównie były to zabezpieczenia typu Awaria. W przypadku zmiany A 60% przypadków zadziałania zabezpieczeń dotyczyło soboty i odpływu F23\_2.

W drugiej kolejności skupiono się na parametrze dotyczącym działania zabezpieczeń na pierwszym odpływie tj. Ozn\_podzesp = F23\_1, który w głównej mierze zasila urządzenia wykorzystywane w robotach eksploatacyjnych. Jako parametry algorytmu wprowadzono minimalne wsparcie na poziomie 0,1, zaufanie na poziomie 0,5 oraz minimalną długość reguł jako 2. Wyznaczone reguły przedstawiono w tabeli 6.

Na podstawie otrzymanych wyników można stwierdzić, że w ponad połowa przypadków zadziałania zabezpieczeń na pierwszym odpływie wystąpiło podczas zmiany D, dotyczyły one zabezpieczenia zwarciowego oraz trwały od 1 do 5 min.

Wyznaczone reguły można zaprezentować również w formie graficznej: wykresu sieciowego (Rys.7a i b) lub macierzowego (Rys.8).

Na wykresach grafowych miary jakości reguł



Tab. 5. Przykłady reguł z poprzednikiem „Zmiana=C” i „Zmiana=A”, źródło: opracowanie własne

Tab. 5. Examples of rules with predecessor „Zmiana=C” and „Zmiana=A”

Reguły	Wsparcie	Zaufanie	Przyrost
{Zmiana=C} => {Dzien_tyg=piątek}	0.364	0.769	2.015
{Zmiana=C} => {Miesiac=2015-01}	0.382	0.808	1.645
{Zmiana=C} => {Nazwa_zab=Awaria}	0.364	0.769	1.567
{Zmiana=A} => {Dzien_tyg=sobota}	0.109	0.6	4.714
{Zmiana=A} => {Ozn_podzesp=F23_2}	0.109	0.6	1.138

Tab. 6. Przykłady reguł z poprzednikiem „Ozn\_podzesp=F23\_2”, źródło: opracowanie własne

Tab. 6. Examples of rules with predecessor „Ozn\_podzesp=F23\_2”

Reguły	Wsparcie	Zaufanie	Przyrost
{Ozn_podzesp=F23_1} => {Zmiana=D}	0.200	0.55	1.891
{Ozn_podzesp=F23_1} => {Nazwa_zab=Zabezpieczenie_zwarciove}	0.218	0.60	1.375
{Ozn_podzesp=F23_1} => {Czas_trwania_dyskr=1 min - 5min}	0.200	0.55	1.210

podane są w postaci etykiet (węzłów i strzałek) oraz kolorów. Wielkość węzła ilustruje wsparcie danego parametru charakteryzującego zadziałanie zabezpieczenia, kolor symbolizuje parametr przyrostu, natomiast strzałki wzajemne zależności [4]. Wykresy te bardzo dobrze sprawdzają się w małych zestawach reguł, w przypadku bardziej skomplikowanych zależności (lub większej liczby parametrów) stają się mniej czytelne.

Innym przykładem wizualizacji reguł asocjacji jest wykres macierzowy. Sposób przedstawienia miary jakości reguł jest analogiczny do wykresów grafowych. Różnica polega na przedstawieniu zależności. W tym przypadku reguły odczytywane są z osi wykresu, poprzednik z osi poziomej, następnik z osi pionowej. Na osi poziomej oprócz poprzednika, dostępne są także informacje o długości poprzednika oraz liczności jego występowania.

#### 5.4. Ocena wyników

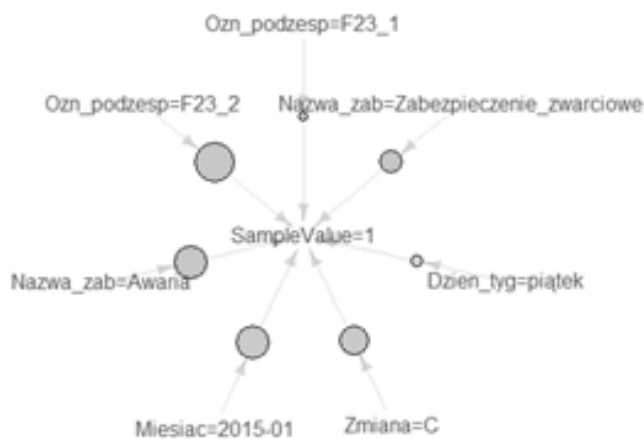
W wyniku przeprowadzonej analizy otrzymano reguły o wysokim i dużym stopniu zaufania. Pozwalają one na wskazanie pojawiających się

zależności warunkujących pracę aparatury zabezpieczeniowej.

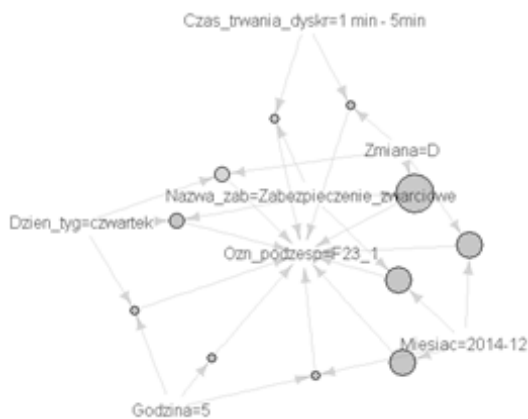
Dla potrzeb osób zarządzających procesem warto byłoby wyjaśnić, co takiego działo się w styczniu, że wystąpiło w nim bardzo dużo przypadków zadziałania zabezpieczeń, szczególnie w odniesieniu do pracy zmiany C. Jak przebiegały prace na zmianie A w soboty i dlaczego najwięcej przypadków zabezpieczenia zwarciowego na pierwszym odpływie dotyczy zmiany D. Taka wiedza pomoże w sformułowaniu poprawnych wniosków dla podjęcia dalszych decyzji w zakresie minimalizacji strat czasu, które mają wpływ na wyniki realizacji procesu w oddziale.

#### Podsumowanie

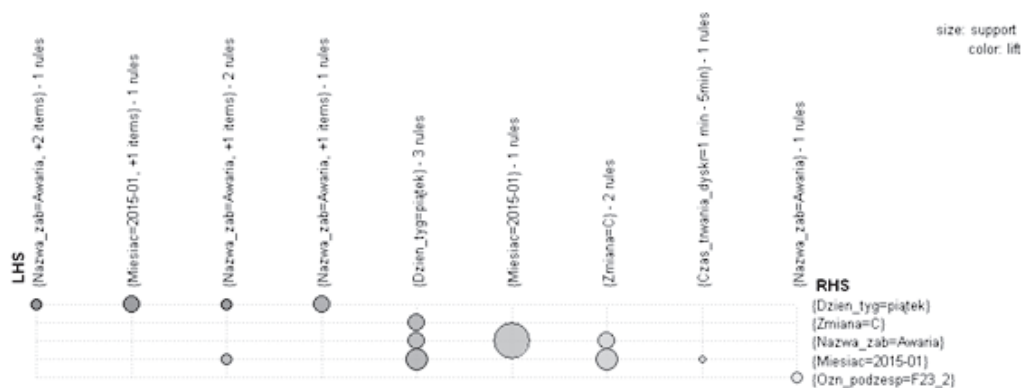
Reguły asocjacyjne są jedną najczęściej wykorzystywanych technik analizy danych do poszukiwania zależności w dużych zbiorach danych. W przedstawionym artykule technika ta została wykorzystana do analizy pracy aparatury zabezpieczeniowej stacji transformatorowej, używanej w oddziałach wydobywczych w kopalniach podziemnych.



Rys.7a. Reguły dotyczące zmiennej „Sample\_Value=1” jako następnik, źródło: opracowanie własne  
 Fig. 7a. Rules for variables "Sample\_Value=1" as consequent



Rys.7b. Reguły dotyczące poprzednika „Ozn\_podzesp=F23\_1”, źródło: opracowanie własne  
 Fig. 7b. Rules for predecessor „Ozn\_podzesp=F23\_1”



Rys. 8. Macierz reguł dotyczących tylko przypadków zadziałania zabezpieczenia, źródło: opracowanie własne  
 Fig. 8. Rule matrix for cases of protection activation

Na podstawie algorytmu Apriori wyznaczono 18 reguł o wysokim i dużym stopniu zaufania. Wstępnie stwierdzono, że czynnikami sprzyjającymi przypadkom zadziałania zabezpieczenia była zmiana C, dzień tygodnia piątek, miesiąc styczeń, zabezpieczenie Awaria oraz odpływ F23\_2. Pogłębiona analiza wykazała, iż wszystkie awarie w styczniu na zmianie C pojawiły się w piątek. W styczniu te wyłączenia w przeważającej części charakteryzowały się czasem trwania 1-5 min. Większość awarii wykazywanych było na odpływie F23\_2. Z reguł wynika, iż że większość przypadków zadziałania zabezpieczeń na zmianie C było w piątek, najwięcej w styczniu i zabezpieczenia typu Awaria. W przypadku zmiany A, większość przypadków zadziałania zabezpieczeń dotyczyło soboty i odpływu F23\_2. Ponad połowa przypadków zadziałania zabezpieczeń na pierwszym odpływie wystąpiło podczas zmiany D, dotyczyły one zabezpieczenia zwarcioowego oraz trwały od 1 do 5 min.

Na podstawie ogólnej analizy pracy stacji transformatorowej można stwierdzić, iż przypadki

zadziałania zabezpieczeń występują bardzo rzadko. Najczęściej są to jednak zdarzenia dotyczące sytuacji potencjalnie niebezpiecznych. Wszelkie informacje umożliwiające, chociaż w niewielkim stopniu, predykcję ich wystąpienia są zatem bardzo cenne. Otrzymane reguły dostarczyły nieoczywistych informacji m.in. wskazały, iż najczęściej zadziałanie zabezpieczeń wystąpiło w warunkach mało intensywnej pracy np. zmiana C oraz na odpływie, który w małym stopniu zasila kompleks ścianowy.

Zaprezentowany w artykule przykład pokazuje możliwości wykorzystania zaawansowanych technik eksploracji danych w analizie przebiegu procesów przemysłowych. Otrzymane wyniki w postaci reguł asocjacyjnych mogą stanowić dla zarządzających wstęp do pogłębionej analizy realizowanego procesu.

*Artykuł opracowano w ramach badań statutowych prowadzonych w Akademii Górniczo-Hutniczej im. St. Staszica w Krakowie (umowa 11.11.100.693).*

## Literatura – References

1. Brzychczy E., Metoda modelowania i optymalizacji robót eksploatacyjnych w wielozakładowym przedsiębiorstwie górniczym, Wydawnictwa AGH, Seria Rozprawy i Monografie, nr 245. Kraków, 2012.
2. Cichosz P., Systemy uczące się, Wyd. WNT, Warszawa, 2007.
3. Hahsler M., Mining Association Rules and Frequent Itemsets, CRAN, 2016.
4. Hahsler M., Chelluboina S., Visualizing Association Rules: Introduction to the R-extension Package arulesViz, CRAN, [przełądane 21.06. 2016]. Dostępne w : <https://cran.r-project.org/web/packages/arulesViz/vignettes/arulesViz.pdf>.
5. Harańczuk G., Przewidywanie awarii i problemów z jakością. StatSoft Polska 2013 [przełądane 21.06.2016]. Dostępne w: [http://www.statsoft.pl/portals/0/Downloads/Przewidywanie\\_awarii\\_i\\_problemov\\_z\\_jakoscia.pdf](http://www.statsoft.pl/portals/0/Downloads/Przewidywanie_awarii_i_problemov_z_jakoscia.pdf)
6. Kęsek M., Język R w odkrywaniu wiedzy górniczej, Przegląd Górniczy, nr 9, 2010
7. Larose T.D., Odkrywanie wiedzy z danych, Wyd. Naukowe PWN, Warszawa, 2006.
8. Ognioszczelna stacja transformatorowa EH-d31-2600/6,0/3,3/2/01(02), Karta produktu, Elgór + Hansen S.A., [przełądany 21.06.2016]. Dostępny w: [http://elgorhansen.com/assets/files/PL/g1/PL\\_EH\\_d31\\_2600\\_60\\_33\\_2\\_01.pdf](http://elgorhansen.com/assets/files/PL/g1/PL_EH_d31_2600_60_33_2_01.pdf)
9. Williams G., J., Rattle: A Data Mining GUI for R, TheR Journal Vol. 1/2, 2009.
10. Yanchang Z., R and Data Mining: Examples and Case Studies, 2015.

### *Association Rules in Analysis of Selected Machine Working in a Mining Division*

*The article presents the application of advanced data mining techniques to analyse the operation of protection device of transformer station used in underground mines. The aim was to find the factors that favour the occurrence of protection activation. In the research association rules were used. Calculations were performed using the R environment and the addition Rattle (Graphical User Interface for Data Mining in R). As a result various rules were obtained, describing dependencies in work of protection devices for selected transformer station.*

*Keywords: Data analysis, mining, association rules, R language, transformer station*