Transactions of the VŠB – Technical University of Ostrava, Mechanical Series
No. 1, 2017, vol. LXIII
article No. 2027

Lenka LANDRYOVA*, Patrik URBAN **

THE USE OF ASSOCIATION RULES IN THE ANALYSIS OF THE TRENDS OF ALARMS

VYUŽITÍ ASOCIAČNÍCH PRAVIDEL PŘI ANALÝZE ALARMOVÝCH TRENDŮ

Abstract

This article deals with acquired alarm logs analyses, the alarm logs from control systems, and explains the reasons for its collecting and implication of the analyses. Several existing platform solutions can be used to analyze the alarms from the historical trends, and solutions, which offer the filtering of events based on time units to obtain the data about the maximum and average number of logs. This paper contributes to this area by developing a graphical interface for a system that makes it possible to use the analysis by association rules and brings the possibility of detecting frequent and repeating patterns in acquired alarm data files.

Abstrakt

Tento článek se zabývá sběrem a analýzami alarmových logů, alarmových hlášení z řídicích systémů, a vysvětluje důvody pro jejich sběr a důsledky analýz. Několik stávajících řešení na softwarových platformách může být k analýze alarmů použito z ukládaných historických trendů a řešení, která nabízejí filtrování událostí na základě časových období a údajů o maximálním a průměrném počtu logů. Tento příspěvek přispívá k této oblasti vytvořením grafického rozhraní pro systém, který umožňuje použití analýzy pomocí asociačních pravidel a přináší možnost detekce častých a opakujících se vzorů v získaných datových souborech s alarmy.

Keywords

Alarm, Analyses, C#, Association Rules, Algorithms, Testing.

1 INTRODUCTION

Automatic control of production processes has long been among the topics of manufacturing companies. This is because they try to minimize the costs associated with the control of production process. The technical means and instrumentation used to measure, visualize and control production process are different, but the automation of operations always minimizes human factor failure. Managing the alarms in control systems of different manufacturers is dealt by using very complex solutions for data analysis, specifically working with data files based on event and alarm identification. Analyzing and identifying alarms is done, briefly speaking, by a time analysis of received data. Such tools and complex solutions are, for example, Delta V (EMERSON, 2015), Alarm Adviser (Alarm Adviser, 2017), or Iconics – Alarm Analytics (ICONICS, 2010).

- DeltaV Analyze enables a complex view at alarms and events recorded in data files and databases and the interface shows on a panel the statistical data of alarms, events, an alarm

* Assoc. Prof. Ing. Lenka Landryova, CSc., Department of Control Systems and Instrumentation, Faculty of Mechanical Engineering, VŠB - Technical University Ostrava, 17.listopadu 15, 708 33 Ostrava - Poruba, Czech Republic, phone: (+420) 597 324112, e-mail: lenka.landryova@vsb.cz.

** Ing. Patrik Urban, Department of Control Systems and Instrumentation, Faculty of Mechanical Engineering, VŠB - Technical University Ostrava, 17.listopadu 15, 708 33 Ostrava - Poruba, Czech Republic, e-mail: patrik.urban@vsb.cz.

system performance based on EEMUA 191 standard, a detailed overview based on a selected time interval, the 10 most frequent alarms, etc.

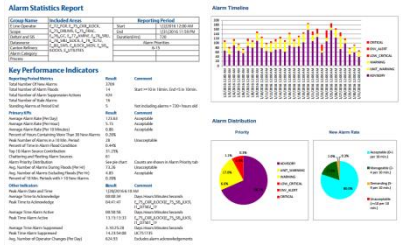

Fig.1: Alarm tools of DeltaV Analyze

- Alarm Adviser is an application interacting with the user to analyze alarms and increase the value of alarming messages in applications, to prevent human errors resulting from overloading due to incorrect operation and alarm configuration. It consists of three parts that can be independently installed on one or more computers due to load distribution, such as Adviser Collector, which collects data, Adviser Service sending and storing data in a database, Adviser Web Server analyzing and publishing calculations on a web site.
- Iconics – Alarm Analytics enables visualization and analyses, alarm management and maintenance according to the best industrial practice. A proper analysis of alarms can reveal significant opportunities to improve the existing operation and mitigation of exceptional situations. Alarm Analytics is built on the BizViz platform, and records and analyzes all information from alarms and events to identify frequent alarms, unacknowledged alarms, and a number of alarm questions. The advantage of Alarm Analytics is in archiving alarms and operators responses into a database, analyzing them in real time according to EEMUA 191 alarm performance standard, and exporting them into user reports.

Fig.2: Alarm Adviser tool

Fig.3: Alarm Analytics tool

2 THE PERFORMANCE OF CONTROL SYSTEMS AND THEIR ALARM ANALYSES

It is generally known that it is possible to control only the system, which is monitored, its performance is measured and the automatic control is communicated to a higher level of control with a human supervision. This hierarchy of control for a control system architecture is standardized in ISA95 standard, which is described in a model for a control system infrastructure. Standard ISA-95 consists of five level descriptions. The first level description contains the standard terminology and an object model, the second part considers object attributes, the third part is aimed on functions and activities of higher levels of control, the fourth part considers object models and the attributes of management of manufacturing operations and the fifth part considers the transactions between business and production, see Fig. 4.

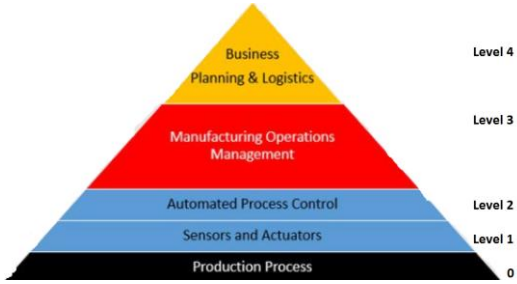

Fig.4: ISA 95 Manufacturing Control Model, where data originates

Such a complex control system can be demonstrated, for example, on a control system of marine technologies. The terminology for "marine" is a well-known and broad concept dealing with control systems for ships, vessels, floating equipment, such as oilrig platforms, etc., see Fig. 5., with possible division according to the type of drive, the propulsion system, the way of navigation and the construction of the ship's body, interior design, voyage target, and operational deployment by purpose. According to these, and from the purpose and operational deployment point of views, the concept can be defined and applied on ocean vessels with personal transport consisting of the following subsystems:

- Vessel Management System (VMS)
- Process Power Management (PPS)
- Heat, Ventilation and Air Conditioning (HVAC)
- Electronic Shut Down (ESD)


Fig.5: Illustrative photos of a ship and oilrig

From these distributed systems at the levels 0-2, the data files stored and communicated to a higher level of control contain attributes of all monitored variables with items of defined levels for each alarm triggered. It is there, where an alarm and event data originate. A trend can be created, which follows the data, where each element complements the sum of received the alarms in a time interval. The basis of such observations is the monitoring of a life cycle for safety and ensuring security, which is a part of the acceptance of functional safety as an independent area of system reliability.

The reliability of the system can be described as the ability of the system to perform the given functions within a given time interval while maintaining the values of the defined operational variables within given limits. The system security integrity results from the rate of occurrence of dangerous malfunctions as defined in the fourth part of IEC 61508 standard as "*the probability that the security system will satisfactorily perform the required safety functions under the given conditions within the specified time*."

2.1 The methodology of data files processing

In the tested file, the data from a sample input file represent a variable‘s values. In the case, when they cannot be sorted out and their comparison cannot be done, they are considered as nominal variables. Then, they can be statistically described as they follow an absolute and relative frequency number, where:

- The absolute frequency number is defined as the number of occurrence of a given variance of a qualitative variable, where *n* represents an element variable of the sample:

$$n_1 + n_2 + \cdots + n_k = \sum_{i=1}^{n} n_i = n \quad (1)$$

- The relative frequency number p_i is defined as a rate of the number of occurrences n_i given to the total number of variables *n*:

$$pi = \frac{ni}{n} * 100\ [\%] \quad (2)$$

2.2 The data sample of the marine data files

The data sample consists of alarm data defined on monitored variables. They are analyzed according to the alarm data standard and the example in Table 1 shows the sum of alarms during one monitored 10 minute interval.

Tab.1: Received alarm frequency, absolute and relative, for the monitored variables’ first 10 minute interval

Monitored variable values	Absolute frequency	Relative frequency
GroupHold	2	0.13
FBFault	1	0.07
FCUnit	1	0.07
ObjectERR	2	0.13
PosError	7	0.47
PosErr	2	0.13
DevErr	0	0.00
DeactFail	0	0.00
Total	15	1.00

The trend representing the alarm data frequency for each 10 minute interval is shown in Figure 6.

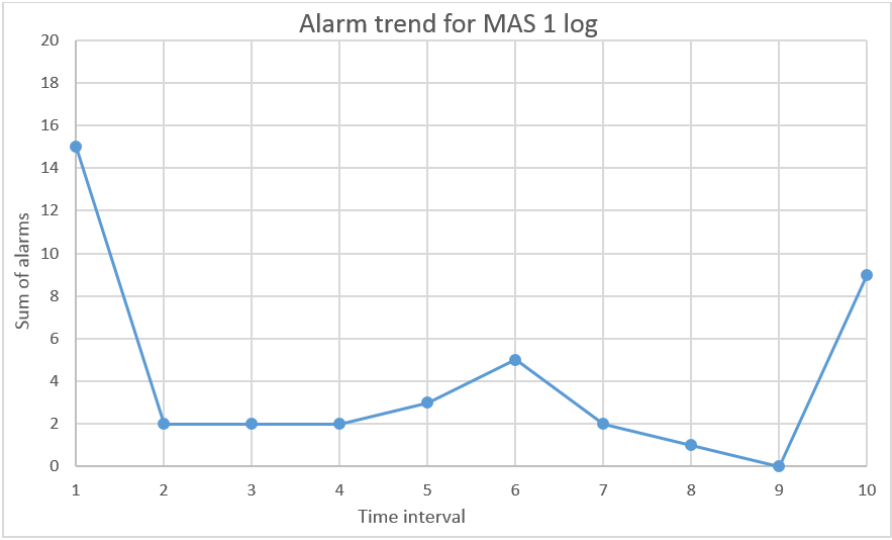

Fig.6: Alarm trend for a marine data log from machinery automated system 1

3 THE USE OF ASSOCIATION RULES

The approach for association rules is according to (Dasseni E., 2001) based on the semantic interpretation and subjective evaluation of the extracted knowledge, where factors, for example support, confidence, presence and absence of expected rules, reverse rules, the relationship between the extracted rules set, and the frequent itemsets, etc. are considered.

At the beginning, the area of knowledge mining was called in many ways: information harvesting, data archeology, data distillery (Briatková, 2008). Knowledge mining and data mining was formalized. Briefly speaking, it is a non-trivial acquisition of "hidden", previously unknown and potentially useful information. Knowledge mining is a process made up of several steps from selecting the necessary data, prepared for analysis, processing, and resulting interpretation. The background of the "frequent pattern" analysis was historically dealing with the research of customers and the products they buy in supermarkets, and their current content in the shopping carts with a certain advance given abundance. The Apriori algorithm is according to (Halaš, 2008) the best-known algorithm for frequent patterns processing.

3.1 The Apriori algorithm on data sets

As mentioned in (Briatková, 2008) and (Halaš, 2008) the Apriori algorithm has the role of finding all of the associations of a set of transactions as well as rules, whose support is equal to or greater than the specified minimum, and the confidence is greater than the minimum entered.

$$supp\ (A) = \frac{number\ of\ transactions\ consisting\ itemset\ A}{number\ of\ all\ transactions}$$

The support is expressing the percentage as number of transactions in a file containing searched items sets.

$$conf\ \left(A \rightarrow B\right) = supp\ (A\ B)/supp(A)$$

Confidence determines the percentage rate of transactions that contain items from A, and at the same time, items from B with definition of the association rule implication (Association Rules, 2010). The algorithm is one of the most used in generating association rules thanks to its simplicity. The

disadvantage of the algorithm is its low efficiency. In each cycle it goes through the entire database and searches for frequent sets. Disadvantages can be eliminated, for example, by HASHing. The hashing table provides a quick search based on pairs index and scrolls through the list.

The main principle of the Apriori algorithm is shown in Figure 7.

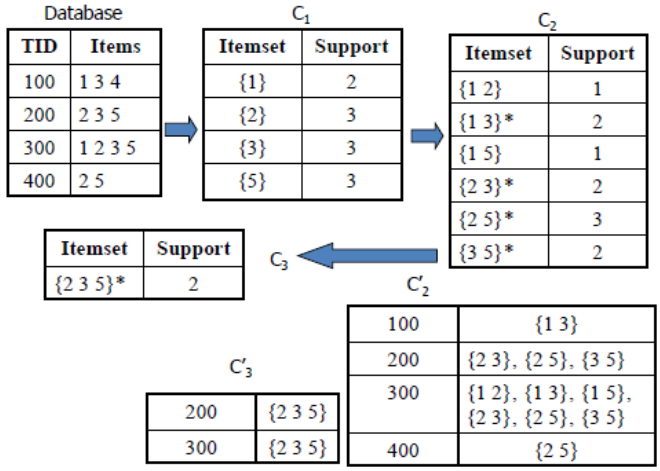

Fig.7: AprioriTid algorithm

3.2 Applying association rules in alarm data trend

The process can be divided into these phases:

1) A list of item sets is generated based on the acquired frequent values of entry items meeting the minSupport condition.

2) New combinations of item sets with the length n + 1 are cyclically generated from the previous item sets, where item sets meeting the minSupport condition are added to the next generation.

3) The item set generation continues until the list of new item sets is equal to 0.

The procedure is described in the following code:

```
public void Run()
{
var ffi = this._ap_object.apriori_math.GET_Candidate(this._ap_object.itemSet);
    this._ap_object.apriori_result.Add(new AprioriResultModel<T>
{ Id = 0, Data = ffi });

    if (this._ap_object.apriori_result.First().Data.Count == 0)
return;
    int counter = 1, count = 0;
    do
    {
var data = ffi.Select(x => x.ItemSet).ToList();
var sfi = this._ap_object.apriori_math.GET_Candidate(data, counter + 1);
count = sfi.Count();
        if (count != 0)
this._ap_object.apriori_result.Add(new AprioriResultModel<T>
{ Id = counter, Data = sfi});
counter++;
      } while (count > 1);
    }}
}
```

The initial step to get a new list of item sets with the length > 1 is to generate of all possible combinations from the item set entry list. The support and confidence values are calculated to the values in the new item set list. The resulting new list of the item set is made up of items that fulfil *support > minSupport* condition. Source code for the list of new itemset is:

```
  public List<AprioriModel<T>> GET_Candidate(List<List<T>> data, int combLength,
bool sort = true, bool eSuppCalc = true, bool eMinSupport = true)
{
var out_data = new List<AprioriModel<T>>();
var work_data = GET_combination(data, combLength);
work_data.ForEach(x => {
 var am = new AprioriModel<T>();
 am.ItemSet = x.ToList();
 am.Support = (GET_support_count_multiple(x.ToList()) /
 _transaction_data.Count) * 100 ;
am.Confidence = (GET_support_count_multiple(x.ToList()) /
 _transaction_data.Select(t => t).Where(t =>
 t.Contains(x.ToList().First())).Count()) * 100;
          out_data.Add(am);
       });
 if (eMinSupport)
 return out_data.Where(x => (x.Support) >= _minSupportValue).ToList();
return out_data;
}
```

The newly developed graphical interface for the pattern search using association rules of the Apriori algorithm allows us to specify the number of time intervals, in which the total list of time intervals should be divided, and the minimum value of minSupport for the support condition. The values are marked in the checkbox as "true".

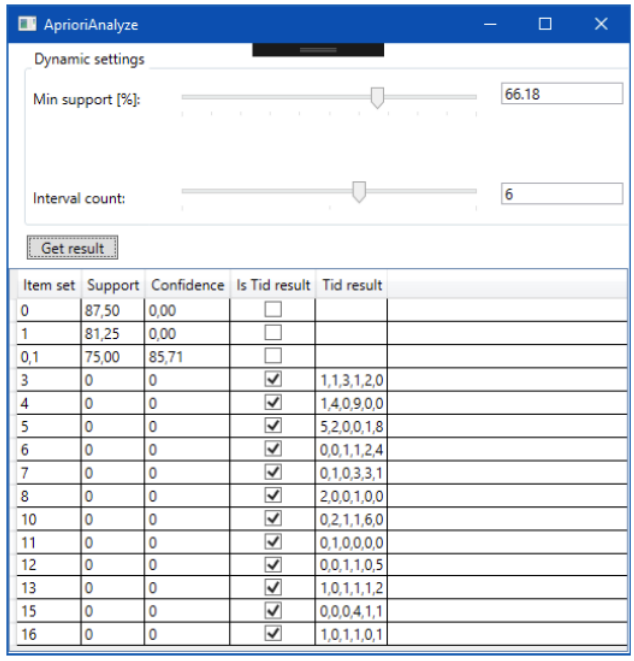

Fig.8: User interface for a search using Apriori association rules

6 CONCLUSIONS

This article discusses the analysis of alarm logs and their trends. Alarm trends give information based on time intervals, when alarm logs are obtained. The time intervals include the elements of logs received. By processing alarm data into the trend the data characteristics show, if they are rising, declining, or do not change. Frequent patterns may occur in alarm log data trends and they may indicate which errors are repeated at certain time intervals. The use of association rules shows an alarm data analysis with the determination of patterns that repeat.

ACKNOWLEDGMENT

The paper has been supported by the project of "Advanced Methods for Process and Machine Control SP2017/106" on the Faculty of Mechanical Engineering, VŠB-TU Ostrava.

REFERENCES

Alarm Adviser. (2017). Retrieved from https://www.wonderware.com/ hmi-scada/alarm-adviser/

Association Rules. An Introduction to Data Mining. (2010). Retrieved from http://www.saedsayad.com/association_rules.htm

Bratkova, M. (2008). *Dobývanie znalostí z e-learningových dát*. Retrieved from https://is.muni.cz/th/72543/fi_m/DP_Briatkova.pdf

Dasseni E., V. V. (2001). *Hiding Association Rules by Using Confidence and Support.* Berlin, Heldelberg: Springer.

EMERSON, P. M. (2015). *DeltaV™ Analyze*. Retrieved from http://www2.emersonprocess.com/siteadmincenter/PM%20DeltaV

ICONICS. (2010). *ICONICS - ALARM ANALYTICS V9.2*. Retrieved from https://www.ertech.ch/servlet/delivered/ProductBriefV9.2_AlarmAnalytics.pdf?magic=cmVzb3VyY2U9NWM4OWFkOC9jYTg4M2JlZS1jZjgwLTRlMWMtOGI4MC0xYjg1NmY4YTg4YTgucGRmJm1pbWU9YXBwbGljYXRpb24vcGRm&name=ca883bee-cf80-4e1c-8b80-1b856f8a88a8.pdf

Kalas, M. (2008). Retrieved from Analýza vybraných metód a algoritmov dolovania v dátach: http://www2.fiit.stuba.sk/~kapustik/ ZS/Clanky0910 /halas/ index.html#_Toc243379371

Wang, H. W. (2008). *A knowledge Management Approach to Data Mining Process for Business Intelligence.* Industrial Management & Data Systems 108(5).

Wang, X. (1999). Data Mining and Knowledge Discovery - an Overview. *Advances in Industrial Control*. London: Springer.