

# Multi-Objective Evolutionary Algorithm for the Identification of Rare Functional Dependencies in Complex Technical Infrastructures

Federico Antonello

*Energy Department, Politecnico di Milano, Via Lambruschini 4, 20156, Milan, Italy.*

Piero Baraldi

*Energy Department, Politecnico di Milano, Via Lambruschini 4, 20156, Milan, Italy.*

*E-mail: [piero.baraldi@polimi.it](mailto:piero.baraldi@polimi.it)*

Enrico Zio

*Energy Department, Politecnico di Milano, Via Lambruschini 4, 20156, Milan, Italy.*

*MINES ParisTech, PSL Research University, CRC, Sophia Antipolis, France.*

*Eminent Scholar, Department of Nuclear Engineering, College of Engineering, Kyung Hee University, Republic of Korea*

Luigi Serio

*Engineering Department, CERN, 1211 Geneva 23, Switzerland*

A Multi-Objective Evolutionary Algorithm (MOEA) is proposed for the identification of association rules describing functional dependencies in Complex Technical Infrastructures (CTIs). The algorithm uses novelty search to explore the solution space. It has been applied to a real large-scale database of alarms collected in the CTI of CERN (European Organization for Nuclear Research). The obtained results show its effectiveness in identifying rare functional dependencies not found using standard algorithms of Association Rule Mining (ARM) algorithms.

*Keywords:* Complex Technical Infrastructures, Dependent Abnormal Behaviours, Alarms, Rare Association Rules, Multi Objective Evolutionary Algorithm; Genetic Algorithm

## 1. Introduction

Functional dependencies in Complex Technical Infrastructures (CTIs) can lead to unexpected cascades of failures (Serio et. al., 2018; Antonello et. al., 2019). Because the architecture of a CTI evolves in time with modifications and extensions to its initial designs, the identification of functional dependencies is a difficult task and classical methods of system decomposition and logic analysis may fail (Zio, 2016; Rebello, et. al., 2018).

To overcome this difficulty, Antonello, et. al. (2019) proposed an Association Rule Mining (ARM) method, which scans alarm databases to identify functional dependencies and groups of functionally dependent components. The method relies on an Apriori-based algorithm that mines alarm databases to extract groups of alarms which frequently occur together and derive the association rules among them. Groups of functionally dependent components are, then, identified by analyzing the generated association

rules. Specifically, Apriori-based algorithms employ a level-wise iterative search mechanism, which scans the whole database for identifying “frequent” patterns, and drives the search for other “frequent” patterns which are supersets of the ones previously identified (Srikant and Agrawal 1996). A pattern is considered only if its frequency of occurrence is larger than a predefined threshold, called minimum support. A main challenge in the application of the Apriori-based algorithms to alarm databases is the difficulty of identifying rare functional dependencies, which are typically unknown and can be actually most relevant for CTI vulnerability (Wang, et al., 2000; Kim and Yun, 2016; Zio, 2016; Antonello, et. al., 2019). This is because their identification would require the use of a small minimum support threshold, which is typically unaffordable computationally (Lin and Tseng, 2006, Wulandari, et. al., 2019). The use of a small minimum support threshold generates a very large number of rules, increasing the complexity and making the set of identified rules hard to analyse (Marin, et. al., 2008; Zhang, et. al.,

*Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference*

*Edited by Piero Baraldi, Francesco Di Maio and Enrico Zio*

*Copyright © ESREL2020-PSAM15 Organizers. Published by Research Publishing, Singapore.*

*ISBN: 978-981-14-8593-0; doi:10.3850/978-981-14-8593-0*

2013). Then, a relatively large minimum support threshold has to be employed, but this comes with the risks of *i*) not identifying rare patterns of alarms and *ii*) extracting association rules which are already evident to the CTI operators (Antonello, et. al., 2019).

Multi-Objective Evolutionary Algorithms (MOEAs) have been recently tailored for mining association rules without the intermediate step of frequent pattern mining and without the need of setting a minimum support threshold (Yan X., et. al., 2009; Mukhopadhyay A., et. al. 2014). MOEAs are meta-heuristic approaches inspired by biological evolution, based on operations such as selection, recombination and mutation. The application of MOEAs for rule mining requires to evolve a population of candidate association rules according to properly defined rule metrics (Mukhopadhyay A., et. al. 2014; Badhon, B., et. al., 2019).

An issue related to the use of MOEAs for rule mining is the tendency of converging on one or a limited set of optimal solutions, even though ARM applications usually require the identification of several interesting and different rules (Martín, D., et. al., 2016). Specifically, CTIs are characterized by various functional dependencies which are relevant for the CTI vulnerability and one is interested in identifying all of them (Antonello et. al., 2019). Thus, during the optimization search it is necessary to find multiple solutions and maintain them in the population, for effectively exploring the solution space and preventing premature convergence to local optima. In this respect, novelty search has been proposed in (Lehman and Stanley, 2011) to exhaustively explore the search space. In detail, novelty search drives the selective pressure to favor diversification in the population by dynamically rewarding the individuals based on their novelty with respect to other individuals, instead of rewarding them considering static fitness objectives (Gomes, J., et al 2017). This approach has been successfully used for evolutionary applications in robotics (Lehman and Stanley, 2011; Gomes et al., 2017), neuroevolution (Conti, et. al., 2018), game content generation (Liapis et al., 2015) and for design optimization in concurrent engineering (Lu. et. al., 2018), showing its ability in finding good solutions faster and more exhaustively than

fitness-based evolution (Gomes et al., 2017). However, to the best of the authors' knowledge, novelty search has not yet been applied to ARM.

In this work, we propose a MOEA which combines novelty search and metrics of rule quality for association rule mining, to identify functional dependencies from alarm data in CTIs. The main advantages with respect to traditional ARM algorithms are:

- It does not require setting minimum support and minimum confidence thresholds.
- It allows extracting functional dependencies, independently from their frequencies of occurrence.
- It allows identifying rare functional dependencies with a reasonable computational effort.

The proposed MOEA is applied to a real large-scale database of alarms generated by the supervision systems of different components of the CTI at CERN.

The remainder of the paper is organized as follows. Section 2 describes the problem setting and the alarm database representation. In Section 3, the proposed MOEA is explained. Section 4 introduces the case study and discusses the obtained results. Finally, Section 5 draws some conclusions.

## 2. Problem setting

We consider a CTI composed by a large number of components,  $N_c \gg 1$ , and we assume to have available a database containing a large number of alarm messages,  $N^{al} \gg 1$ , generated by a CTI during a long period of time  $[t_o, t_f]$ . The generic *i*-th alarm message is associated to the pair  $(t_i, m_i)$ , defined by the time  $t_i$  at which the alarm occurs and a label  $m_i$  identifying the type of alarm. Assuming that there are  $M_j^{al}$  different types of alarms associated to the generic *j*-th component,  $c_j$ , we use the label  $a_j^k$  to refer to the *k*-th type of alarm message associated to component  $c_j$ . The set containing all the possible types of alarm messages in the database is:

$$A = \{a_1^1, \dots, a_1^{M_1^{al}}, \dots, a_{N_c}^1, \dots, a_{N_c}^{M_{N_c}^{al}}\} \quad (1)$$

and the total number of alarm message types:

$$M^{al} = \sum_{j=1}^{N_c} M_j^{al} \quad (2)$$

In this work, we consider two components, functionally dependent if the operation of one is influenced by the operation of the other (Estesami et. al. 2016). In particular, we consider alarm messages triggered when components have abnormal behaviours or malfunctions. We assume that a functional dependency among two components of a CTI,  $c_1$  and  $c_2$ , exist if a malfunction of component  $c_1$ , revealed by an alarm,  $a_1^k$ , causes a malfunction of components  $c_2$ , revealed by another alarm,  $a_2^k$ , or viceversa (Antonello et. al., 2019).

### 2.1 Alarm database representation

The entire time period  $[t_0, t_f]$  during which the alarm messages of the database have been collected is subdivided into  $Z$  consecutive small time intervals of the same length  $\Delta t = \frac{t_f - t_0}{Z}$ . A Boolean variable,  $s_j^k(z)$ , is associated to the occurrence of the alarm of type  $a_j^k$  in the  $z$ -th time interval:

$$s_j^k(z) = \begin{cases} 1 & \text{if alarm } a_j^k \text{ occurs at least once in} \\ & [t_0 + (z-1) \cdot \Delta t, t_0 + z \cdot \Delta t] \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

The state of the generic component  $c_j$  in the generic  $z$ -th time interval is represented by the Boolean vector:

$$\vec{c}_j(z) = [s_j^1(z), s_j^2(z), \dots, s_j^{M_j^{al}}(z)] \in [0,1]^{M_j^{al}} \quad (4)$$

and that of the CTI by the concatenation of the component state vectors  $\vec{c}_j(z)$ :

$$\vec{J}(z) = [\vec{c}_1(z), \dots, \vec{c}_{N_c}(z)] \in [0,1]^{M^{al}} \quad (5)$$

Finally, the raw database of alarms  $(t_i, m_i)$ ,  $i = 1, \dots, N^{al}$ , is transformed into the Boolean matrix:

$$T = \begin{bmatrix} \vec{J}(1) \\ \dots \\ \vec{J}(Z) \end{bmatrix} \in [0,1]^{Z \times M^{al}} \quad (6)$$

whose generic  $z$ -th row represents the state of the CTI during the  $z$ -th time interval. Therefore,  $T$  provides a dynamic representation of the CTI state evolution in the time interval  $[t_0, t_f]$ .

### 3. Method

In this Section we describe the encoding-decoding procedure adopted for representing association rules by means of chromosomes of the MOEA (Section 3.1), the novelty search-based MOEA itself (Section 3.2) and the search objectives, initial population and genetic operators (Section 3.3).

#### 3.1 Encoding-Decoding Procedure

Considering a set of alarms  $X \subseteq A$ , an association rule is a probabilistic logical expression of the form  $x^a \Rightarrow y^a$ ,  $x^a \subset X$ ,  $y^a = X - x^a$ , where  $x^a$  and  $y^a$  are referred to as ‘‘antecedent’’ and ‘‘consequent’’ of the rule, respectively (Srikant and Agrawal, 1996). In the proposed approach, we represent a set of association rules using a population of  $N^{pop}$  chromosomes in which each individual chromosome,  $ind_i$ ,  $i = 1, \dots, N^{pop}$ , encodes a rule (Del Jesus, et. el., 2011; Mukhopadhyay, et. al., 2014; Badhon, et. al., 2019). The generic alarm  $a_j^k$  is associated to two genes of the chromosome and each gene contains one bit. If the two bits associated to alarm  $a_j^k$  are  $[1,0]$ , then the alarm  $a_j^k$  belongs to the antecedent of the rule, i.e.,  $a_j^k \subset x^a$ , if the two bits are  $[0,1]$ , then the alarm belongs to the consequent, i.e.,  $a_j^k \subset y^a$ , otherwise, the alarm does not belong to the rule (Del Jesus, et. el., 2011; Mukhopadhyay, et. al., 2014; Badhon, et. al., 2019). Notice that in this way a chromosome is a vector of  $M^{al}$  bits representing an association rule. Figure 1 gives an example of the decoding of a generic binary vector.

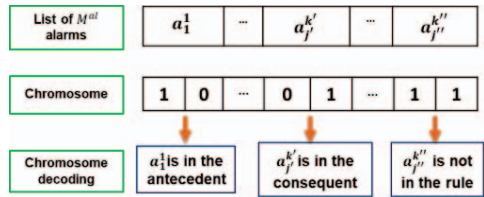


Figure 1. Example of chromosome decoding.

#### 3.2. Novelty Search

Novelty search creates a constant evolutionary pressure to preserve unique and novel individuals

(Lehman and Stanley, 2011; Gomes et al., 2017). Instead of rewarding performance based on fitness objectives, the key idea is to reward divergences from prior behaviors (Lehman and Stanley, 2011). The uniqueness of each individual with respect to the rest of the population is evaluated by characterizing the *sparseness* of the individual in the search space. The areas with denser clusters of chromosomes are considered less novel. Specifically, the *sparseness* of an individual is computed as its average distance to the other individuals in the population. In this work we evaluate the novelty of each individual  $ind_i$  as:

$$Novelty(ind_i) = \frac{1}{N^{pop}} \sum_{j \neq i} dist(ind_i, ind_j) \quad (7)$$

where  $dist$  is a domain-dependent measure of the distance among individuals. The Jaccard distance  $dist_j$ , which has been proved to be effective in evaluating *sparseness* for pattern mining and association rule mining applications (Tummala, et. al. 2018), is here used to evaluate the distance among individuals  $ind_i$  and  $ind_j$ :

$$dist_j(ind_i, ind_j) = dist_j(r_i, r_j) = \frac{|x^a_i \cup x^a_j| + |y^a_i \cup y^a_j| - |x^a_i \cap x^a_j| - |y^a_i \cap y^a_j|}{|x^a_i \cup x^a_j| + |y^a_i \cup y^a_j|} \quad (8)$$

where  $| \cdot |$  refers to the cardinality of the alarm set, i.e. the number of alarms involved in the set. Notice that if  $ind_i$  and  $ind_j$  encode two identical rules, their distance is equal to 0, otherwise, if they encode two rules which do not share any common element neither in the antecedent nor in the consequent part of the rule, their distance is equal to 1.

### 3.3. Evolutionary Algorithm

We use a Genetic Algorithm (GA) due to its straightforward principles, its simplicity of implementation and the fact that it has been already successfully applied to ARM (Anand, et., al, 2009; Badhon, et. al., 2019). As in (Gomes et al., 2017), we combine novelty search and traditional fitness objectives in the NSGA-II MOEA (Deb, 2000), which is considered the most effective optimization algorithm for multi-objective rule mining. The objectives of the MOEA search are the maximization of the novelty measure defined by Eq. 9 and the metrics of rule quality.

With respect to the latter, MOEAs for ARM have been initially proposed to identify association rules based on search mechanisms driven by the metrics of *support* and *confidence* (Del Jesus, et. el., 2011; Mukhopadhyay, et. al., 2014; Badhon, et. al., 2019). Later, additional metrics have been introduced as objectives for the search, such as *interestiness*, *lift*, *comprehensibility*, *cosine*, *prevalence*, *recall*, *Laplace* (Srinivasan and Ramakrishnan, 2001; Del Jesus, et. el., 2011). In this work we consider the metric of *interestiness*, which is the most effective for identifying rare rules (Dhaenens and Jourdan, 2016).

Let  $n(X)$  be the counter of the number of vectors  $\vec{T}(z)$  of the database  $T$  characterized by the occurrences of at least all the alarms of the set  $X$  (i.e.,  $\forall a_{j,s}^k \subset X, s_j^k(z) = 1$ ). *Interestiness* is defined as:

$$Interestiness(x^a \Rightarrow y^a) = \frac{n(x^a \cup y^a)}{n(x^a)} * \frac{n(x^a \cup y^a)}{n(y^a)} * \left(1 - \frac{n(x^a \cup y^a)}{z}\right) \quad (9)$$

Notice that *interestiness* is proportional to 1) how much the consequent is dependent to the antecedent, 2) how much the antecedent is dependent to the consequent, and 3) how much rare is the rule. According to this metric, the most interesting rules are rare, since they are the most likely to be unknown, and characterized by a strong mutual dependency between antecedent and consequent.

Given that the above-mentioned metric tends to favor rules involving low number of attributes (Pachón Álvarez, V., et. al., 2012), in this work we consider as a further search objective the number of alarms involved in the association rules:

$$Length(x^a \Rightarrow y^a) = |x^a \cup y^a| \quad (10)$$

The maximization of the rule *length* allows avoiding the identification of rules made by partial combinations of the alarms: for example, given a functional dependency involving 4 alarms ( $a_{j_1}^{k_1}, a_{j_2}^{k_2}, a_{j_3}^{k_3}$  and  $a_{j_4}^{k_4}$ ), the rules made by 2 alarms out of 4 ( $(a_{j_1}^{k_1}) \Rightarrow (a_{j_3}^{k_3}), (a_{j_2}^{k_2}) \Rightarrow (a_{j_4}^{k_4})$ ) are characterized by large *interestiness*, but are less relevant for the search of functional dependencies than the rules made by 4 alarms.

In (Shenoy et al. 2005), the effectiveness of using

an initial population of rules with small *Length* which satisfies some conditions (e.g., support larger than the minimum threshold) is shown (Del Jesus, M.J, et. el., 2011). In this work, the initial population of chromosomes is created considering all the possible rules characterized by a length equal to 2, and, then, selecting the best  $N^{pop}$  individuals following the NSGA II algorithm.

As suggested by (Del Jesus, et. el., 2011; Mukhopadhyay et. al., 2014) in case of standard binary chromosomes, standard genetic operators and an evolution algorithm based on a traditional genetic algorithm with two-point crossover and flip mutation are used. Furthermore, we avoid the presence of identical individuals in the population to favor population diversity.

#### 4. Case Study: CERN Complex Technical Infrastructure

The CTI of CERN LHC, which is the largest existing particle accelerator in the World, is composed by several systems working together to support the operation of the LHC (Nielsen and Serio, 2016). It consists of a 27 km ring of superconducting magnets and infrastructures, extending over the Swiss and French borders and located about 100 m underground.

In this case study, we consider a database of alarms generated during the period  $[t_0, t_f]$ =[January 1<sup>st</sup>, 2016; December 31<sup>st</sup>, 2016] by three supervision systems of a representative part of the LHC infrastructure, the LHC *point 8*. During the considered period,  $N_{al}$  =18711737 alarms reporting  $M^{al}$ =13451 different types of malfunctions have occurred. The above dataset has been pre-processed by pruning those alarms which are involved in already known functional dependencies among components belonging to a same system. Considering that a functionally dependent group made by  $R$  different alarms potentially generates  $3^R - 2^{R+1} + 1$  different rules (Del Jesus et. al., 2011), and that groups of dependent components belonging to a same system may involve more than 70-80 alarms, the dataset pre-processing avoids the analysis of tens of thousands of non-interesting rules. The pruned dataset consists in  $N_{al}$  =112591 alarm messages reporting  $M^{al}$ =1024 different types of malfunctions.

Considering the expected propagation time of a perturbation due to the system intrinsic physical characteristics (distance, thermal inertia, electrical propagation, fluid flows, etc.), the time interval length is set to  $\Delta t = 30$  min (Antonello et. al., 2019) in order to ensure the identification of all possible rules while minimising computational resources and spurious rules identification. Therefore, the one-year period [January 1<sup>st</sup>, 2016, December 31<sup>st</sup>, 2016] is divided into  $Z = 17500$  time intervals. An initial population of  $N^{pop} = 500$  individuals, encoded into chromosome of  $2 * M^{al}$  bits, is used, the mutation probability is set equal to  $1 / (2 * M^{al})$  and the crossover probability to 0.8 (R. Anand, et. al., 2009). The population is evolved for 2000 generations obtaining a final set of 500 association rules in a computational time of 830 seconds on an Intel core (TM) i7-4790 CPU@ 3.6 GHz, 16 GB RAM. Table 10 reports some examples of extracted rules. The first rule describes the correlation among malfunctions in a breaker cryo safety electric system distribution switchboard (EKD202\_SLASH\_8U) and malfunctions of two pumps in the cooling and ventilation system (SU\_8\_UPKA802\_AL6, SU\_8\_UPEA802\_AL6). The second rule describes the associated occurrence of a malfunction in a breaker of the electrical system (EKD204\_SLASH\_8U) and three different malfunctions in low-pressure compressors or the cryogenic system (QSCB\_8\_CSY\_C2). Rule 3, and Rule 4 describe the propagation of a malfunction triggered by problems in the cryo safety electric system distribution switchboard ('EKD104\_SLASH\_8HM\_I1314', 'EKD107\_SLASH\_8HM\_I1314'), which propagate and lead to malfunctions of the Cryogenic system ('QSAB\_8\_QSA\_TS3.IST', 'QSCB\_8\_CSY\_C1\_S13.IST', 'QSRB\_8\_CV003\_FS1.IST').

N	Antecedent	Consequent
1	EKD202_SLASH_8U	'SU_8_UPEA802_AL6' 'SU_8_UPKA802_AL6'
2	'EKD106_SLASH_8HM, 'QSCB_8_CSY_C3_S13'	'QSCB_8_CSY_C3_S15' 'QSCB_8_3CV120AL'
3	'QSAB_8_QSA_TS3.IST', 'QSRB_8_PV279_S11.IST', 'EKD107_SLASH_8HM_11 314', 'EKD104_SLASH_8HM_11 314'	QSCB_8_CSY_C1_S13.IST', 'QSRB_8_PV100_S11.IST', 'QSRB_8_CV003_FS1.IST', 'SU_8_AC_AL2.IST', 'SU_8_AC_CCC2.IST'
4	'QSRB_8_PV100_S11.IST', 'QSRB_8_PV279_S11.IST', 'SU_8_AC_AL2.IST', 'EKD104_SLASH_8HM_11 314'	'QSAB_8_QSA_TS3.IST', 'QSCB_8_CSY_C1_S13.IST', 'QSRB_8_CV003_FS1.IST', 'SU_8_AC_CCC2.IST', 'EKD107_SLASH_8HM_11314'

Table 1. Example of obtained rules associating alarms generated by components of different systems

An independent expert analysis has confirmed that the involved components are, indeed, part of a chain of malfunctions occurred in 2016.

A detailed analysis of the generated association rules has shown that 53 different functionally dependent groups of alarms have been identified by the proposed method. This confirms the capability of the novelty search to generate a diversified set of results.

The effectiveness of the novelty search is investigated by comparing the results obtained applying MOEA with different combinations of search objectives. The results are shown in Table 2. As expected, when novelty search is not used, the final population converges to a population of rules describing only few functional dependencies.

Table 2. Number of groups of functionally dependent alarms identified considering various combinations of search objectives.

Search objectives	Groups of functionally dependent alarms
Interestingness, Length and Novelty	53
Interestingness and Length	5
Interestingness	4

To further analyse the advantages of the proposed evolutionary approach, a traditional Apriori-based algorithm (Antonello, et. al, 2019) is applied to the same simulated dataset. In order to

be able to identify rare rules, the value of *minimum support* threshold is set equal to 3 and the *minimum confidence* is set equal to 0.6%. The search has produced 1049 association rules in a computational time of 43959 seconds on an Intel core (TM) i7-4790 CPU@ 3.6 GHz, 16 GB RAM.

Notice that the proposed evolutionary approach allows reducing the computational effort (830 seconds) with respect to traditional ARM techniques (43959 seconds), and without requiring the setting of the minimum support. Furthermore, the use of novelty search and rule length reduces the generation of many rules involving the same set of alarms, with respect to traditional ARM. The traditional ARM approach, that mines all the rules which satisfy given support and confidence thresholds, tends to find more rules, some of which are not relevant for functional dependencies identification.

### 5. Conclusions

A MOEA based on novelty search has been developed for the identification of rare functional dependencies in CTIs. An application to a real dataset of the CTI of CERN has shown its ability in exploring the solution space by identifying a larger number of functional dependencies than other MOEAs which do not use novelty search. Furthermore, its comparison with a traditional ARM algorithm shows its capability of identifying rare functional dependencies without requiring the setting of the minimum support, and a reduction in the computational effort required to generate the rules.

### References

Agrawal, R., Imieliński, T., Swami, A. (1993). “Mining Association Rules Between Sets of Items in Large Databases”, ACM SIGMOD Record, Volume 22, Issue 2, 6 January 1993, Pages 207-216.

Antonello, F., Baraldi, P, Shokry, A, Zio, E, Gentile, U., Serio, L. (2019). Data-Driven Extraction of Association Rules of Dependent Abnormal Behaviour Groups. *Proceedings of the 29th International European Safety and Reliability Conference*, ESREL 2019.

Antonello, F., Baraldi, P, Shokry, A, Zio, E, Gentile, U., Serio, L. (2019). Association Rules Extraction for the Identification of Functional Dependencies in Complex Technical Infrastructures. *Reliability Engineering & System Safety* Under review, (Submitted 2019).

- Etesami, J., Kiyavash, N. (2017), "Measuring causal relationships in dynamical systems through recovery of functional dependencies" (2017) *IEEE Transactions on Signal and Information Processing over Networks*, 3 (4), art. no. 7782866, pp. 650-659.
- Filip, F. G. (2008). "Decision Support and Control for Large-Scale Complex Systems." *Annual Reviews in Control* 32(1): 61–70.
- Serio, L., Antonello, F., Baraldi, P., Castellano, A., Gentile, U., Zio, E. (2018), "Smart framework for the availability and reliability assessment and management of accelerators technical facilities" 9th International Particle Accelerator Conference, IPAC 2018
- Wang, X. Analysis of the time complexity of quick sort algorithm (2011). Proceedings 4th International Conference on Information Management, Innovation Management and Industrial Engineering, ICIII 2011, 1, art. no. 6115063, pp. 408-410.
- Zhang, J., Wang, Y., Feng, J (2013). Attribute index and uniform design based multiobjective association rule mining with evolutionary algorithm (2013) *The Scientific World Journal*, 2013, art. no. 259347.
- Marin, N., Molina, C., Serrano, J. M. and Vila, M. A. (2008). A complexity guided algorithm for association rule extraction on fuzzy datacubes. *IEEE Transactions on Fuzzy Systems*, vol. 16, no. 3, pp. 693–714, 2008.
- Mukhopadhyay A., Maulik U., Bandyopadhyay S., Coello C.A.C., (2014). Survey of multiobjective evolutionary algorithms for data mining: Part II. (2014) *IEEE Transactions on Evolutionary Computation*, 18 (1), art. no. 6658840, pp. 25-35.
- Rebello, Yu, Ma, (2018). An integrated approach for system functional reliability assessment using Dynamic Bayesian Network and Hidden Markov Model", *Reliability Engineering & System Safety*, Volume 180, 2018, Pages 124-135.
- Srikant, Ramakrishnan, and Rakesh Agrawal. (1996). "Mining Quantitative Association Rules in Large Relational Tables." *ACM SIGMOD Record* 25(2): 1–12.
- Lin, W. Y., & Tseng, M. C. (2006). Automated support specification for efficient mining of interesting association rules. *Journal of Information Science*, 32(3), 238–250.
- Kim, D., & Yun, U. (2016). Efficient mining of high utility pattern with considering of rarity and length. *Applied Intelligence*, 45(1), 152–173.
- Badhon, B., Kabir, M.M.J., Xu, S., Kabir, M. A (2019). Survey on association rule mining based on evolutionary algorithms. *International Journal of Computers and Applications*
- X. Yan, C. Zhang, S. Zhang (2009). Genetic algorithm-based strategy for identifying association rules without specifying actual minimum support. *Expert Systems with Applications* 36 (2009) 3066–3076.
- Shenoy PD, Srinivasa KG, Venugopal KR, Patnaik LM (2005). Dynamic association rule mining using genetic algorithms. *Intelligent Data Analysis*, 9(5):439–453, 2005.
- Del Jesus, M.J., Gámez, J.A., González, P., Puerta, J.M (2011). On the discovery of association rules by means of evolutionary algorithms 2011 *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 1(5), pp. 397-415
- Anand, R., Vaid, A., Singh, P.K (2009). Association rule mining using multi-objective evolutionary algorithms: Strengths and challenges. 2009 *World Congress on Nature and Biologically Inspired Computing, NABIC 2009 – Proceedings* 5393878, pp. 385-390
- Srinivasan and Ramakrishnan. (2011). Evolutionary multiobjective optimization for rule mining: A review. *Artif. Intell. Rev.*, vol. 36, no. 3, pp. 205–248, Oct. 2011.
- Pachón Álvarez, V., Mata Vázquez, J. (2012). An evolutionary algorithm to discover quantitative association rules from huge databases without the need for an a priori discretization (2012) *Expert Systems with Applications* 39(1), pp. 585-593
- Deb, K., Agrawal, S., Pratap, A., Meyarivan, T. (2000). A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 1917, pp. 849-858
- Martín, D., Alcalá-Fdez, J., Rosete, A., Herrera, F. (2016). NICGAR: A Niching Genetic Algorithm to mine a diverse set of interesting quantitative association rules. *Information Sciences* 355-356, pp. 208-228
- Gomes, J., Mariano, P., Christensen, A.L. (2017). Novelty-driven cooperative coevolution. *Evolutionary Computation* 25(2), pp. 275-307
- Conti, E., Madhavan, V., Such, F.P., Lehman, J., Stanley, K.O., Clune, J. (2018). Improving exploration in evolution strategies for deep reinforcement learning via a population of novelty-seeking agents. *Advances in Neural Information Processing Systems*, 2018-December, pp. 5027-5038.
- Liapis, A., Yannakakis, G. N., and Togelius, J. (2015). Constrained novelty search: A study on game content generation. *Evolutionary Computation*, 23(1): 101–129

- Dhaenens, C., Jourdan, L. *Metaheuristics for Big Data* (2016) *Metaheuristics for Big Data*, 5, pp. 1-188.
- Tummala, K., Oswald, C., Sivaselvan, B. (2018). A Frequent and Rare Itemset Mining Approach to Transaction Clustering. *Communications in Computer and Information Science* 804, pp. 8-18.
- Wulandari, C.P., Ou-Yang, C., Wang, H.-C (2019). Applying mutual information for discretization to support the discovery of rare-unusual association rule in cerebrovascular examination dataset (2019) *Expert Systems with Applications*, 118, pp. 52-64.
- Zio, (2016). Challenges in the vulnerability and risk analysis of critical infrastructures, *Reliability Engineering & System Safety*, Volume 152, 2016, Pages 137-150, ISSN 0951-8320.