Bayesian Network Representation of Meaningful Patterns in Electricity Distribution Grids

Hassan M. Nemati, Anita Sant'Anna, Sławomir Nowaczyk
Center for Applied Intelligent Systems Research (CAISR)
Halmstad University
Halmstad, Sweden
hassan.nemati@hh.se, anita.santanna@hh.se, slawomir.nowaczyk@hh.se

***Abstract*—The diversity of components in electricity distribution grids makes it impossible, or at least very expensive, to deploy monitoring and fault diagnostics to every individual element. Therefore, power distribution companies are looking for cheap and reliable approaches that can help them to estimate the condition of their assets and to predict the when and where the faults may occur.**

In this paper we propose a simplified representation of failure patterns within historical faults database, which facilitates visualization of association rules using Bayesian Networks. Our approach is based on exploring the failure history and detecting correlations between different features available in those records. We show that a small subset of the most interesting rules is enough to obtain a good and sufficiently accurate approximation of the original dataset. A Bayesian Network created from those rules can serve as an easy to understand visualization of the most relevant failure patterns. In addition, by varying the threshold values of support and confidence that we consider interesting, we are able to control the tradeoff between accuracy of the model and its complexity in an intuitive way.

***Index Terms*—Smart Grids, Condition Monitoring, Data Mining, Failure Statistics, Association Rules, Bayesian Networks.**

I. Introduction

Industries, infrastructure, and citizens are relying on power electricity, and consequently electricity outages can have disastrous effects on them. This raises a need for proper strategies for power failure prediction and prevention. One of the common approaches for this purpose is exploiting failures history. In this paper we present three different directions for detecting fault patterns in an electricity distribution grid: *Failure Statistics*, *Association Rules*, and *Bayesian Networks*. The results can be used to design better maintenance strategies to prevent common outages.

Failure statistics is a practical technique for component reliability evaluation which is addressed in many previous works [6]–[8], [10], [11], [13]. In this case, failure probability, failure rate, principal failure causes, events classification, and mean-time-between-failures (MTBF) are the most commonly considered indicators.

In general, the utility companies keep records for previous faults that contain features describing the event. To analyze failures characteristic it is crucial to discover which failures have common features, e.g., if there are any types of failures that happen mostly in certain parts of the network or at certain times. In the literature this analysis is known as discovering patterns in event sequences. One approach is based on finding association rules, and has been addressed by several researchers [1], [3], [5], [14]. Association rules are based on the frequency of the co-occurrence of features and conditional dependency between them. Their interestingness is often expressed in terms of probability.

Bayesian Networks [2], [9], [12] are graphical representations of probabilistic relationships over a set of variables, constructed using probability distribution over a set of variables in a dataset. If we consider features of failure events as probabilistic variables, a Bayesian Network captures the conditional relations between those features over a set of events. In [4], Fauré C. et al. used a five step algorithm to model the frequent association rules using Bayesian Networks. The first step of their algorithm is to create a Bayesian Network based on expert domain knowledge, and then compute association rules for all the combinations of features. The interestingness measure of a rule, derived from the knowledge-driven Bayesian Network, makes it possible to filter out unimportant rules. Tian D. et al. in [12] proposed a Bayesian association rule mining algorithm (BAR) which combines the association rule algorithm with Bayesian Networks. They compute association rules for all the combinations of features and then construct a Bayesian Network. Finally they interpret the result using Bayesian confidence and Bayesian lift using the interestingness of the association rules.

Our approach is based on Tian's result, however, we do not need to compute the association rules for all combinations of features, which is a very time consuming task. Instead, we construct the Bayesian Network by using only the rules with two features. We show that such a network is still a good approximation of the original dataset, and most of the strong associations can be represented by the joint probability distributions.

In this paper, we explore the real fault history of an electricity distribution company from south of Sweden. Therefore, first we present a simple statistical analysis of historical failures in this grid. Then we present analysis of relations

978-1-4673-8463-6/16/$31.00 © 2016 IEEE

between features (time, place, the corresponding main-station and sub-station, switchgear, voltage level, cause of the failure, etc.) in the dataset and their co-occurrences. We compute association rules between features, and select the ones with high support and confidence as interesting, since they are describing high statistical correlations. Finally, we propose a Bayesian Network representation of the association rules by using only the rules with two features. We show that most of the strong association rules are conveniently represented by the joint probability distributions of this Bayesian Network, by making some assumptions about the dependency between directed and undirected nodes. This representation provides a simplified visualization of the conditional connections between features and at the same time is much faster to compute than the association rules for all combinations of features.

II. Meaningful Pattern Discovery

A. Failure Statistics

Primary evaluation of the historical failure is used for analyzing the frequency of occurrence for each failure in an electricity distribution grid. In this case two factors are commonly considered: the probability of occurrence, and the mean-time-between-failures (MTBF). The first factor is calculated as the ratio of the number of each failure over the total number of failures during a specific time interval, and can be used for determining e.g. the most unreliable components. The second factor is the expected number of days between failures of a given type.

B. Mining Association Rules

The objective of mining association rules is to find the most frequently occurring combinations of features. Let $I = \{I_1, I_2, ..., I_m\}$ be a set of features (items). An association rule is an implication of the form $A \Rightarrow B$, where $A \subset I$, $B \subset I$, and A, B are disjoint itemsets, i.e. $A \cap B = \emptyset$. In this case the itemset $A = \{a_1, a_2, ...\}$ is the prior and the itemset $B = \{b_1, b_2, ...\}$ is the posterior of the rule. Now assume that $X = \{x_1, x_2, ..., x_n\}$ is a set of random variables representing the list of observations (events) in a dataset. Each observation x_i in the dataset X may or may not contain a specific item e.g. $x_1 = \{I_1, I_2, I_5\}$ only contains items I_1, I_2, I_5.

We define X_A as $X_A = \{x_i \in X$ that contains items $A\}$. In this case, the *support* of itemset A, represented by $S(A)$, is the ratio of the cardinality of X_A over the cardinality of the dataset X

$$S(A) = \frac{|X_A|}{|X|} = P(X_A). \tag{1}$$

The support of a rule, denoted as $S(A \Rightarrow B)$, is the percentage of observations in the dataset that contain both A and B:

$$S(A \Longrightarrow B) = \frac{|X_{A\cup B}|}{|X|} = \frac{|X_A \cap X_B|}{|X|} = P(X_A, X_B). \tag{2}$$

The *confidence* of an association rule is the percentage of examples containing A that also contain B; or, in other words, a fraction that shows how frequently B occurs among all the observations containing A:

$$C(A \Rightarrow B) = \frac{S(A \Rightarrow B)}{S(A)} = P(X_B|X_A) \tag{3}$$

The confidence value indicates how reliable the rule is and in this work we use it to measure the interestingness of the pattern.

The *lift* of an association rule is a ratio of the confidence of the rule to the frequency of observations containing B. It is a value between 0 and infinity that measures the deviation of a rule from statistical independence:

$$L(A \Rightarrow B) = \frac{C(A \Rightarrow B)}{S(B)} = \frac{P(X_A, X_B)}{P(X_A)P(X_B)}. \tag{4}$$

A lift value smaller than one indicates negative correlation, equal to one indicates no correlation, and greater than one indicate positive correlation between features A and B among all the observations.

C. Constructing Bayesian Networks

The confidence value of each association rule corresponds to the strength of the conditional dependence between features, therefore, they can be used for automatically building a Bayesian Network.

An association rule $A \Rightarrow B$ can be seen as a connection from one itemset to another. If $I = \{I_1, I_2, ..., I_t, ..., I_m\}$ is a set of features such that $A = \{I_1, I_2, ..., I_t\}$ and $B = \{I_{t+1}, ..., I_m\}$, the Bayesian Network representation for all the connections between feature set A and B is shown in Figure 1, where items in set A are parent nodes and items in set B are child nodes.

In general, the joint probability distribution represented by a network can be written as:

$$P(I_1, I_2, ..., I_t, ..., I_m) = \prod_{i=1}^{m} P(I_i|parents(I_i))$$

In our case, where the itemset A is the parent of itemset B, the joint probability distribution represented by the network can be written as:

$$P(I_1, I_2, ..., I_t, ..., I_m) = \prod_{i=1}^{t} P(I_i) \prod_{j=t+1}^{m} P(I_j|I_1, I_2, ..., I_t) \tag{5}$$

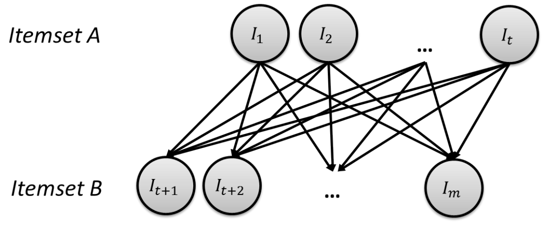

Fig. 1. The Bayesian Network representing association rule $(A \Rightarrow B)$

As shown in equation (2), each of the terms $P(I_j|I_1, I_2, ..., I_t)$ corresponds to the confidence of the rule $((I_1, I_2, ..., I_t) \Rightarrow (I_j))$.

D. Reasoning with Bayesian Networks

We would like to compute the conditional probability of more than two items e.g. $P(X_A|X_B, X_C)$ by using the characteristics of Bayesian Networks and the available connections between set of two items. For this purpose we consider two situations:

If there is no direct connection between X_B and X_C in the Bayesian Network, we make the simplifying assumption of their independence, which allows us to use formula (6) to approximate $P(X_A|X_B, X_C)$ by the following:

$$\begin{aligned} P(X_A|X_B, X_C) &= \frac{P(X_A, X_B, X_C)}{P(X_B, X_C)} = \\ &= \frac{P(X_B, X_C|X_A)P(X_A)}{P(X_B)P(X_C)} = \\ &= \frac{P(X_B|X_A)P(X_C|X_A)P(X_A)}{P(X_B)P(X_C)} = \\ &= \frac{P(X_A|X_B)P(X_A|X_C)}{P(X_A)} \end{aligned} \tag{6}$$

This value can be calculated directly from the Bayesian Network, without the need to reference the original data.

If, on the other hand, there exists at least one direct connection in the Bayesian Network between X_B and X_C, the assumption of their independence would lead to too significant errors. In this case we need to consider the two nodes to be dependent. We define the connection with highest confidence as the "primary" connection (without loss of generality we assume that it is from X_C to X_B) and approximate $P(X_A|X_B, X_C)$ by the following:

$$\begin{aligned} P(X_A|X_B, X_C) &= \frac{P(X_A, X_B, X_C)}{P(X_B, X_C)} = \\ &= \frac{P(X_A, X_B, X_C)}{P(X_B|X_C)P(X_C)} \end{aligned} \tag{7}$$

In the following section we show that these simplifying assumptions still provide, in practice, good enough approximations of the actual empirical probabilities.

III. Experimental Results

In this study we consider the Halmstad Energi och Miljö electricity distribution grid (HEM Nät) in the south of Sweden. We use failure history in the grid as our input dataset. This dataset contains information about historical failures for the entire grid during years 2009 to 2015, with a total of 1110 failures. Information such as date, time, the corresponding main- and sub-station, cause of the failure, the faulty component, and outages duration for all the failures in the grid are registered as features.

To discover fault patterns we represent the result using three different methods: failure statistics, association rules, and Bayesian Networks.

A. Failure Statistics

The failures in our dataset are grouped into operational, environmental, and unknown failures. The operational failures are caused by a defect in an internal component, the environmental failures are caused by an external factors, and the root cause of failures for unknown faults are unspecified. In Table I the frequency of different types of operational and environmental failures (in percent) and the MTBF from 2009 until the end of 2014 are calculated. According to this table the most common failures during these years are "Fabrication fault" and "Fuse break" with frequency 34.59% and 24.95%, respectively. On the other hand, the most common environmental failure during this period is "Digging" which happened 14.41% of the time. According to the MTBF, the "Fabrication fault" and "Fuse break" occurred in average every 5.7 days and 7.91 days respectively. The failures caused by "Digging" have occurred every 13.69 days.

TABLE I
Failure statistics of cause of failure in historical data from 2009 until 2015

Type of Failure	Cause of Failure	Frequency(%)	MTBF(days)
Operational Failure (846 failures)	Fabrication fault	34.59%	5.7
	Fuse break	24.95%	7.91
	Incorrect installation	7.12%	27.72
	Overload	5.59%	35.32
	Incorrect operation	1.35%	136.88
	Lack of maintenance	1.44%	146
	Others	1.17%	168.46
Non-Operational Failure (227 failures)	Digging	14.41%	13.69
	Traffic	1.71%	115.26
	Weather	3.42%	57.63
	Animal	0.72%	273.75
	Others	0.18%	1095

TABLE II
Failure statistics of affected component in historical data from 2009 until 2015

Type of Failure	Affected Component	Frequency(%)	MTBF(days)
All Type of Failure (1110 failures)	Underground cable pillar	48.11%	4.1
	Underground feeder cable	26.94%	7.32
	Underground cable fuse	10.09%	19.55
	Concr.sec.substation indoor man	4.32%	45.63
	OH uninsulated free line	2.70%	73
	Others	7.84%	25.17

In general, the primary root cause of "Fabrication fault" is aging of components. This reasoning is also supported by Table II where the "Underground cable pillar" and the "Underground feeder cable" are the components with the most faults during the period, with frequency of 48.11% and 26.94% respectively. Therefore, one of the important issues that should be considered by the company is updating the repair strategies and maintenance scheduling to discover the weak components such as underground feeder cables and underground cable pillar.

B. Association Rules

Each recorded fault is characterized by different features describing the failure. In order to discover association rules between these features, we first create a boolean representation of those features. Then association rule mining techniques are used to discover the frequency of co-occurrence of the features and their correlation. In the experiments we have used the following:

- **Season** — Spring (Mar, Apr, May), Summer (Jun, Jul, Aug), Autumn (Sep, Oct, Nov), Winter (Dec, Jan, Feb)
- **Weekday**
- **Hour** — Hour morning (7-12), Hour lunch (12-13), Hour afternoon (13-18), Hour evening (18-22), Hour night (23-7)
- **Main-station** — H2, H3, H4, H7, H8, H10 and Others
- **Outage duration** — less than one hour ($T<1$), between one and two hours ($1<T<2$), between two and three hours ($2<T<3$), between three and four hours ($3<T<4$), greater than 4 hours ($T>4$)
- **Sub-station** — there are 199 sub-stations in the network
- **Cause of failure** — there are 27 types of failures, of which the most important ones are listed in Table I
- **Affected component** — there are 21 types of components, of which the most important ones are listed in Table II

The association rules with high support and confidence, which also have lift greater than 1, are considered as interesting rules. Selected rules are shown in Table III. These rules can be interpreted in the following ways:

row number 1 — the probability of an outage of duration less than one hour occurring when the affected component is an underground cable pillar is 23.514%. These two are the most common items happening together.

row number 4 — the probability that an underground cable pillar is the affected component knowing that the cause of failure is fuse break is 79.061%.

row number 14 — whenever there is a failure in sub-station N78 it is summer, and the probability of seeing this failure again is 0.36%. These two items are highly correlated since the lift is 3.437.

row number 19 — if there is thunder and it is summer, there is a probability of 0.811% that a failure occurs at main-station H7.

row number 21 — knowing that there is a digging in areas connected to main station H7 that affects the underground feeder cable, we can expect to have long duration outage (between 2-3 hours) in part of the grid with probability 3.153%.

Some of these rules confirm intuitions our expectations regarding certain types of failures. For example, the rule in row 11 confirms the fact that digging would cause failure in underground feeder cables. Another example is $Overload \Rightarrow Winter$, with frequency 39 and confidence 62.5%: if we know that the cause of failure is overload, then its a high probability that the season is winter. Similarly, the rule $Sunday, Hourmorning \Rightarrow 2<T<3$ occurs 32 times with confidence 66.672%. It states that if a failure happens on Sunday morning, it takes a long time to be discovered and repaired.

TABLE III
INTERESTING RULES FOR HISTORICAL FAILURES DURING YEARS 2009 TO 2015

	Prior	Posterior	Frequency	Support	Confidence	Lift
1	T<1	G_Cable Pil	261	23.514	56.739	1.179
2	H7	2<T<3	251	22.613	54.329	1.119
3	Fabrication	2<T<3	229	20.631	59.635	1.228
4	Fuse break	G_Cable Pil	219	19.730	79.061	1.643
5	G_Feeder Ca	2<T<3	172	15.495	57.525	1.185
6	Hour_aftern	2<T<3	163	14.685	50.938	1.049
7	H3	G_Cable Pil	153	13.784	54.064	1.124
8	Winter	G_Cable Pil	150	13.514	55.556	1.155
9	Spring	2<T<3	141	12.703	52.809	1.088
10	Hour_evenin	2<T<3	116	10.450	51.327	1.057
11	Digging	G_Feeder Ca	103	9.279	64.375	2.390
12	N8	Fabrication	4	0.360	100.000	2.891
13	N341	Fabrication	4	0.360	100.000	2.891
14	N78	Summer	4	0.360	100.000	3.437
15	Fuse break, T<1	G_Cable Pil	147	13.243	82.584	1.717
16	Digging, G_Feeder Ca	2<T<3	65	5.856	63.109	1.300
17	Digging, H7	2<T<3	54	4.865	65.064	1.340
18	Saturday, G_Cable Pil	2<T<3	37	3.333	52.116	1.073
19	Thunder, Summer	H7	9	0.811	81.817	1.966
20	H3, G_Cable Pil, T<1	Fuse break	48	4.324	60.002	2.404
21	Digging, H7, G_Feeder Ca	2<T<3	35	3.153	66.035	1.360
22	Overload, Winter, T<1	G_Cable Pil	12	1.081	50.004	1.039
23	Digging, Hour_mornin, G_Feeder Ca , H7	2<T<3	15	1.351	57.701	2.142

C. Bayesian Networks

We propose a new approach where association rules with two items are used for constructing a Bayesian Network. For this purpose the lists of priors and posteriors of each rule correspond to the network nodes, and the connections between them correspond to the conditional dependency between random variables. In this section we show that the Bayesian Network constructed based on the interesting rules of two items is a good approximation of the real dataset and it can be used for calculating conditional probabilities of association rules for more than two items. To this end we compute the conditional probabilities of association rules for three features using several examples. Four classes of relations between features and the corresponding connections are shown in Figure 2.

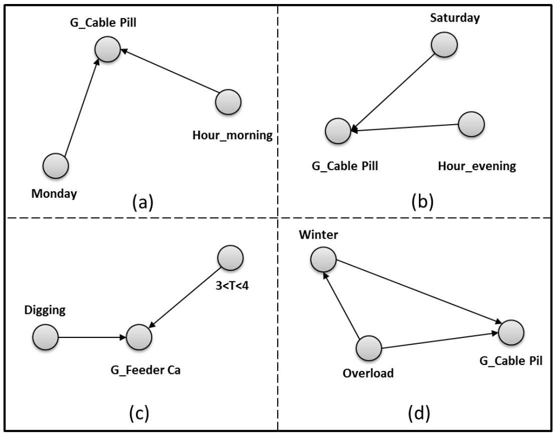

Fig. 2. Some examples of connected nodes selected from Bayesian Network

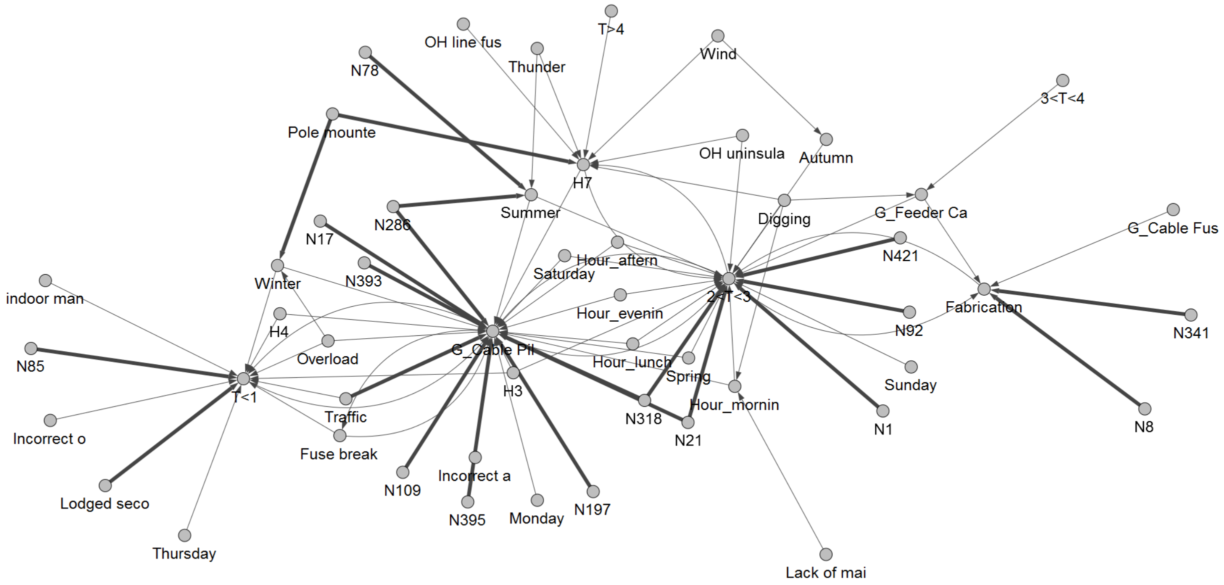

Fig. 3. Bayesian Network representation of association rules constructed from strong connection between two features

The fully connected failure network contains $2^{n=276}$ connections between all the nodes. However, we assume that some of the items are independent or the dependencies can be neglected since they are very weak (confidence smaller than a certain threshold). Considering this assumption the constructed Bayesian Network for our dataset is shown in Figure 3. In this figure, the threshold values are set such that itemsets with frequency higher than or equal to 200 have at least 40% confidence level, itemsets with frequency between 100 to 200 have at least 45% confidence level, itemsets with frequency between 10 to 100 have at least 50% confidence level, itemsets with frequency between 5 to 10 have 80% confidence level, and itemsets with frequency equal to 4 have 100% confidence level. The connections with confidence greater than 80% are shown with bold arrows which corresponds to the availability of strong confidence (more than 80%) between the two nods.

Based on this network and formula (6) we can calculate the conditional probability of three features when two of them are independent. Examples of such features are (a) and (b) in Figure 2:

$$P(G\_CablePil|Monday, Hourmorning) \\ = \frac{P(G\_CablePil|Monday).P(G\_CablePil|Hourmorning)}{G\_CablePil} \\ = 50.56\%$$

$$P(G\_CablePil|Saturday, Hourevening) \\ = \frac{P(G\_CablePil|Saturday).P(G\_CablePil|Hourevening)}{P(G\_CablePil)} \\ = 61.404\%$$

Based on this network and formula (7) we can calculate the conditional probability of three features which have direct connections and thus are not independent. Examples of such features are (c) and (d) in Figure 2:

$$P(3{<}T{<}4|G\_FeederCa, Digging) \\ = \frac{P(3{<}T{<}4, G\_FeederCa, Digging)}{P(G\_FeederCa, Digging)} \\ = \frac{P(3{<}T{<}4, G\_FeederCa, Digging)}{P(G\_FeederCa|Digging).P(Digging)} = 18.429\%$$

$$P(G\_CablePil|Overload, Winter) \\ = \frac{P(G\_CablePil, Overload, Winter)}{P(Overload, Winter)} \\ = \frac{P(G\_CablePil, Overload, Winter)}{P(Winter|Overload).P(Overload)} = 51.284\%$$

Although the joint probability value of feature sets with more than three features can be calculated directly from dataset, the selection of interesting itemset and their conditional probability are captured from the important and strong connectivity between nodes in the Bayesian Network (Figure 3).

The comparison between results of computing the above conditional probabilities from the association rules with three features and the Bayesian Network are shown in Table IV.

It can be seen that, except for the probability of $P(G\_CablePil|Saturday, Hourevening)$, the probabilities computed from the Bayesian Network are very close to what we computed directly from the dataset. The difference between results for the probability of $P(G\_CablePil|Saturday, Hourevening)$ comes from the

TABLE IV
COMPARISON BETWEEN RESULTS USING THE ASSOCIATION RULES AND THE BAYESIAN NETWORK

Conditional probability example	Result from the Bayesian Network	 the dataset
P(G_Cable Pil$\|Monday, Hourmorning$)	50.56%	53.806%
P(G_Cable Pil$\|Saturday, Hourevening$)	61.404%	76.649%
P(3<T<4$\|G_FeederCa, Digging$)	18.429%	12.62%
P(G_Cable Pil$\|Overload, Winter$)	51.284%	51.275%

first assumption, i.e., that when there is no direct connection between two nodes we assume they are independent.

Note that, these conditional probabilities can also be directly calculated from the dataset but representing the connectivity between nodes using Bayesian Network makes it easier to pick features corresponding to the meaningful patterns.

Another representation example the Bayesian Network constructed from association rules is shown in Figure 4. In this network the threshold values are set such that itemsets with frequency higher than or equal to 10 have at least 50% confidence level, itemsets with frequency between 5 to 10 have 80% confidence level, and itemsets with frequency equal to 4 have 100% confidence level.

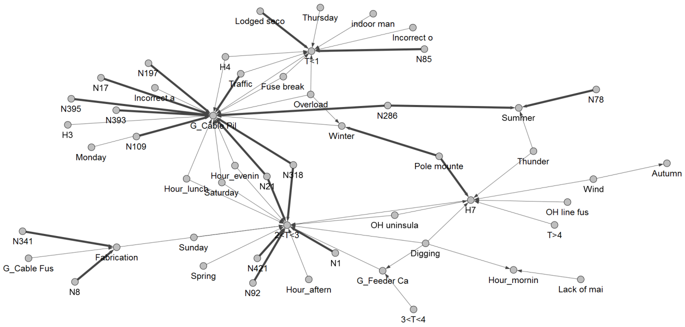

Fig. 4. Statistical failures with Bayesian Network representation using higher threshold value for confidence compare to Figure 3

For features where both of the networks in Figure 3 and Figure 4 have the same connections, the joint probability distribution is equal. However, some joint probabilities such as $P(Sunday, Hourmorning, 2{<}T{<}3)$ is captured by direct connections in Figure 3 while the direct connection between $Hourmorning$ and $2{<}T{<}3$ is not available in Figure 4, therefore the result will be different.

Overall, by varying the threshold values of support and confidence, we are able to control the tradeoff between accuracy of the model and its complexity.

IV. CONCLUSION AND DISCUSSION

In this paper, we give an example of the use of statistical analysis, association rules, and Bayesian Networks to explore the fault history of an electricity distribution grid in the south of Sweden. We present a simple statistical analysis of historical failures looking at the probability of occurrence and the MTBF for different failure causes and components. Then, we present analysis of relations between failure features, in particular their co-occurrences using association rules.

Finally, we show that the association rules can be used to construct a Bayesian Network that can be applied as an intuitive visualization of the conditional relations between features. In this case, most of the strong association rules can be represented by the joint probability distributions of the Bayesian Network while using only the rules with two features. Those results provide a clear and practical representation of features associated with events that can be used by managers and maintenance staff at the company.

REFERENCES

[1] Rakesh Agrawal and Ramakrishnan Srikant. Mining sequential patterns. In *Data Engineering, 1995. Proceedings of the Eleventh International Conference on*, pages 3–14. IEEE, 1995.

[2] Christopher M Bishop. *Pattern recognition and machine learning*. springer, 2006.

[3] Thomas G Dietterich and Ryszard S Michalski. Discovering patterns in sequences of events. *Artificial Intelligence*, 25(2):187–232, 1985.

[4] Clément Fauré, Sylvie Delprat, Jean-François Boulicaut, and Alain Mille. Iterative bayesian network implementation by using annotated association rules. In *Managing Knowledge in a World of Networks*, pages 326–333. Springer, 2006.

[5] Jiawei Han and M Kamber. Mining frequent patterns, associations, and correlations. *Data Mining: Concepts and Techniques (2nd ed., pp. 227-283). San Francisco, USA: Morgan Kaufmann Publishers*, 2006.

[6] Hassan M Nemati, Anita Sant'Anna, and Slawomir Nowaczyk. Reliability evaluation of underground power cables with probabilistic models. In *Proceedings of the International Conference on Data Mining (DMIN)*, page 37. The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp), 2015.

[7] OA Quiroga, J Meléndez, and S Herraiz. Fault causes analysis in feeders of power distribution networks. In *International Conference in Renewables Energies and Quality Power, ICREP*, volume 11, 2011.

[8] Li Qun. Analysis on fault statistics in wenzhou electric power distribution network. In *Power and Energy Engineering Conference, 2009. APPEEC 2009. Asia-Pacific*, pages 1–4. IEEE, 2009.

[9] Stuart Russell and Peter Norvig. Artificial intelligence: a modern approach. 1995.

[10] D Saxena, K Verma, and S Singh. Power quality event classification: an overview and key issues. *International Journal of Engineering, Science and Technology*, 2(3):186–199, 2010.

[11] Thomas Allen Short. *Electric power distribution handbook*. CRC press, 2014.

[12] David Tian, Ann Gledson, Andreas Antoniades, Aristo Aristodimou, Ntalaperas Dimitrios, Ratnesh Sahay, Jianxin Pan, Stavros Stivaros, Goran Nenadic, Xiao-jun Zeng, et al. A bayesian association rule mining algorithm. In *Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference on*, pages 3258–3264. IEEE, 2013.

[13] Xiang Zhang and Ernst Gockenbach. Component reliability modeling of distribution systems based on the evaluation of failure statistics. *Dielectrics and Electrical Insulation, IEEE Transactions on*, 14(5):1183–1191, 2007.

[14] Gao Zhanjun, Peng Zhengliang, Gao Nuo, and Chen Bin. A distribution network fault data analysis method based on association rule mining. In *Power and Energy Engineering Conference (APPEEC), 2014 IEEE PES Asia-Pacific*, pages 1–5. IEEE, 2014.