Mobile Networks and Applications (2019) 24:1713–1721
https://doi.org/10.1007/s11036-018-1047-6

Incorporating FAIR into Bayesian Network for Numerical Assessment of Loss Event Frequencies of Smart Grid Cyber Threats

Anhtuan Le[1] · Yue Chen[1] · Kok Keong Chai[1] · Alexandr Vasenev[2] · Lorena Montoya[2]

Published online: 1 September 2018

Abstract

In today's cyber world, assessing security threats before implementing smart grids is essential to identify and mitigate the risks. Loss Event Frequency (LEF) is a concept provided by the well-known Factor Analysis of Information Risk (FAIR) framework to assess and categorize the cyber threats into five classes, based on their severity. As the number of threats is increasing, it is possible that many threats might fall under the same LEF category, but FAIR cannot provide any further mechanism to rank them. In this paper, we propose a method to incorporate the FAIR's LEF into Bayesian Network (BN) to derive the numerical assessments to rank the threat severity. The BN probabilistic relations are inferred from the FAIR look-up tables to reflect and conserve the FAIR appraisal. Our approach extends FAIR functionality by providing a more detailed ranking, allowing fuzzy inputs, enabling the illustration of input-output relations, and identifying the most influential element of a threat to improve the effectiveness of countermeasure investment. Such improvements are demonstrated by applying the method to assess cyber threats in a smart grid robustness research project (IRENE).

Keywords Cyber threat · Loss event frequency · Threat assessment · Risk management

1 Introduction

Smart grids are recognized as one of the most important elements for the power industry to sustain energy utilization [1]. With its advance abilities in sensing, communication, and actuation, the next decade will see a considerable development in power system operations. These developments will result, among others, in improved real-time information gathering and processing, sophisticated demand-response solutions, and advanced system management solutions [2].

Due to its digitalization, smart grids are facing great challenges from cyber-attacks. Recent examples have shown how such attacks can significantly affect power systems and human life [3]. Risk from cyber threats is therefore a critical issue to consider when implementing smart grids.

Proper risk management comprises of risk assessment, its translation into impact assessment and risk mitigation via countermeasures. In addition, it is important to note that risk acceptance is an important component of risk management, as it acknowledges that not all risk can or should be mitigated. The identification of cyber threats is a crucial component of smart grid risk assessment. However, this task is challenging because of the grid complexity. In addition to a number of potential vulnerabilities, the introduction of new functions and components may lead to new cyber attack vectors [4], hence the risk assessment process must be a sustained effort.

Recently, a number of studies have focused on smart grid threat assessment, especially from the lens of information security. One can usually apply quantitative, qualitative, and hybrid approaches. The ultimate goal of the quantitative

✉ Anhtuan Le
A.Le@qmul.ac.uk; a.le.1@warwick.ac.uk

Yue Chen
Yue.Chen@qmul.ac.uk

Kok Keong Chai
Michael.Chai@qmul.ac.uk

Alexandr Vasenev
A.Vaseneva@utwente.nl

Lorena Montoya
L.Montoya@utwente.nl

[1] Queen Mary University of London, Mile End, E1 4NS, London, UK

[2] University of Twente, Drienerlolaan 5, 7522 NB Enschede, Netherlands

approach is to utilize probability theory and statistics to assign numerical probabilistic values to threat likelihood [5]. While these methods can provide clear guidance about the threat, they have high difficulty in implementation and ambiguity evaluation [6]. On the other hand, the qualitative techniques provide a systematic expert analysis to give a qualitative output rather than a numerical result [7]. Their main advantage is the reliable reasoning; however, in many cases, the output is not detailed enough to take clear decisions [8]. Besides, several hybrid models were proposed that combine the quantitative and qualitative methods and eliminate their weaknesses.

Factor Analysis of Information Risk (FAIR) framework [9] is well-known among hybrid approach. Due to its effective yet simple guidelines for practice, it is applicable in many risk and threat assessment situations. For instance, FAIR is also used to analyze smart grid threats by the Bell Labs Advisory Service [10].

FAIR operates with probabilities and provides a taxonomy of factors contributing to risk and how they affect each other. The risk is defined as "the probable frequency and probable magnitude of future loss" [9]. By acknowledging that risk is an uncertain event, FAIR points out that one should not focus on what is possible, but on how probable a given event is. Thus, this approach focusses on establishing accurate probabilities for the frequency and magnitude of loss events. Even though risk is seen as an uncertain event, it is not clear how to establish the loss event frequency accurately if an input parameter is described as a distribution. In addition, a critical issue relates to establishing how uncertainties in the input parameters propagate and whether this affects the final results. In other words, the final user i.e. the risk manager, should be aware of the stability of the risk assessment results. This is crucial for the next step i.e. the mitigation plan, which is drawn after the identification of countermeasures for each risk and the assessment of their implementation costs versus their risk reduction benefits.

In contrast to a pure qualitative analysis, FAIR assesses threat through the Loss Event Frequency (LEF) concept with a five-point scale, which is suitable for dealing with a small number of threats. However, in smart grids, the number of cyber-threats is often large and will increase following the discovery of new attacks. As a result, there might be many threats falling into the same category. For example, even though the IRENE project [11] has refined a list of 102 threats (outlined in NIST 800−30 [12]) into a list of 38 potential smart grid cyber-threats, the analysis of these threats using FAIR can be difficult as FAIR provides little guidance on how to handle threats within the same class. Thus, finding a way to rank threats within a class, with respect to the input uncertainty, would be desirable. This can be of particular use when an analyst needs to construct an argument for providing sufficient security when resources are limited.

In this paper, we propose a method to incorporate the FAIR framework into Bayesian Network (BN) to obtain numerical threat assessments. BN is a strong tool commonly used in reasoning structural analysis frameworks similar to FAIR. We infer the BN probabilistic relations from FAIR tables to reflect and conserve the FAIR appraisal. Our method provides several following advantages:

- *Supports ranking threats in the same group:* By providing a numerical answer to system managers, we aim to support their perception of threat LEF with respect to other threats in the group. In this case, the managers can make better decisions regarding security countermeasures and mitigation plans;
- *Allows to obtain answers even with fuzzy inputs:* for instance, when experts do not fully agree on specific threat parameters;
- *Illustrate input-output relation:* Illustrates how changes in the input data propagate through the network and contribute to the output;
- *Points out the most influential factor* that, if lowered, can decrease the overall LEF quicker than others.

The remainder of this paper is organized as follows. In Section 2, we describe our model to transform the FAIR framework to the Bayesian network reasoning. Section 3 presents the experiment results and relevant discussions, and Section 4 concludes the paper.

2 Proposed solution

2.1 The FAIR framework

The Loss Event Frequency (LEF) component of the risk can be assessed by adopting the FAIR taxonomy by structurally reasoning about a number of threats factors. These factors include:

- Contact (C): the chance that the attackers, who have access to the asset, will act against it;
- Action (A): the motivation of the attackers when attacking, in particular the type and severity of the attack impact to the asset;
- Threat Capability (Tcap): the probability of the attackers to overcome the protective system; and
- Control Strength (CS): the compensating controls implemented to prevent the potential attacks.

The reasoning structure between these factors are given in Fig. 1. A more detailed tutorial for assessing threat by FAIR is given in [13].

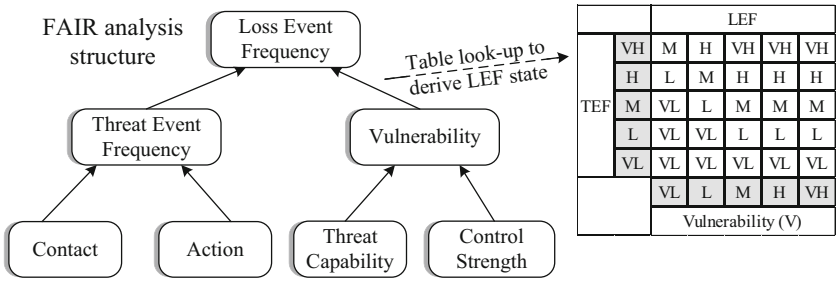

Fig. 1 The reasoning structure of the LEF in the FAIR model

FAIR can support encoding each threat factor by means of a five-point scale (i. e. [Very Low, Low, Moderate, High, and Very High] or [VL L M H VH]). FAIR provides a reference for estimating the state of the input factors. A more detailed way of deriving the input state using the FAIR reference can be found in [13]. Based on the states of the causes, FAIR provides the reasoning tables to look up the state of the effect. The look-up table in Fig. 1 gives an example of how the LEF factor can be derived from the Threat Event Frequency (TEF) and the Vulnerability (V) factors. If values for C, A, CS, and Tcap are provided, the state of the TEF and V can be derived, and leading to the state of LEF.

The FAIR approach is closely related to qualitative definitions of risk listed in [14]. These definitions suggest that risk can be considered as: an exposure to a proposition (e.g., the occurrence of a loss) of which one is uncertain; the occurrences of some specified consequences of the activity and associated uncertainties; and the effect of uncertainty on objectives. The latter definition is provided by ISO and can be seen as a causal relation between uncertainties and objectives. This highlights the importance of analyzing uncertainties to better understand risks.

As with any other model, uncertainty propagates within it to its output. A number of aspects/factors/causes of uncertainty include: lack of information (or knowledge); approximation; abundance of information (or knowledge); conflicting nature of information/data; measurement error; linguistic ambiguity (e.g., the meaning of word "large"); and subjectivity of analyst judgement [15]. Analyzing how such uncertainty propagates within a model is therefore of interest. FAIR as a model provides a guidance for this. The impact of input uncertainty can be assessed if both aleatory and epistemic uncertainties are encoded as input parameters, propagated through the model, and studied in relation to each other at the output. This allows an analyst to understand the relations between input parameters and ultimately about how to quantify the degree of belief, and ultimately how well the real context is represented. Employing well-developed probabilistic analysis for modelling risks and its elements is therefore an important research direction.

2.2 Incorporating a structural analysis into Bayesian network

Bayesian Network (BN) is a strong probabilistic tool used widely to assess the uncertainty based on observed evidences. The BN model consists of nodes representing the variables for reasoning, and edges connecting the nodes to indicate their relations. The probabilistic reasoning is embodied into the conditional probability table (CPT) in each node.

A BN model has the form of a structural analysis with the cause-effect relations between the nodes. On the other hand, querying the probability of a BN structural reasoning can be done once the probabilistic relations (CPT) of every node are provided. CPTs are normally obtained based on evidence; however, they can also be estimated via other methods when the training data is unavailable. In this section, we summarize the fuzzy comprehensive evaluation method from [16], which provides a way to construct the CPTs of a BN model based on the fuzzy evaluation of its cause-effect relations. This method allows the incorporation of a structural analysis into a BN, which will be applied later to transform FAIR to BN.

The method in [16] is as follows. Assume a Bayesian reasoning structure with n **causes** that lead to an **effect** as can be seen in Fig. 2. The causes and the effect all can have m states *1, 2, ..., m*. The impact of each state of cause i to effect is represented through the *fuzzy judging vector* $[r_{i1}r_{i2}...r_{im}]$, with $\sum_{j=1}^{m} r_{ij} = 1, i = 1, 2, ..., n$. Besides, the correlation between the causes is represented through the *weight vector* $[a_1, a_2, ..., a_n]$ with $\sum_{i=1}^{n} a_i = 1$.

The CPTs are obtained via the following formula:

$$P(E = j|C_1 = j_1, C_2 = j_2, ..., C_n = j_n) = \sum_{i=1}^{n} a_i r_{i\sigma(j_i - j, j)} \quad (1)$$

in which $P(E = j|C_1 = j_1, C_2 = j_2, ..., C_n = j_n)$ is the conditional probability of the event in which the effect E has state j, while its corresponding causes $C_1, C_2, ..., C_n$

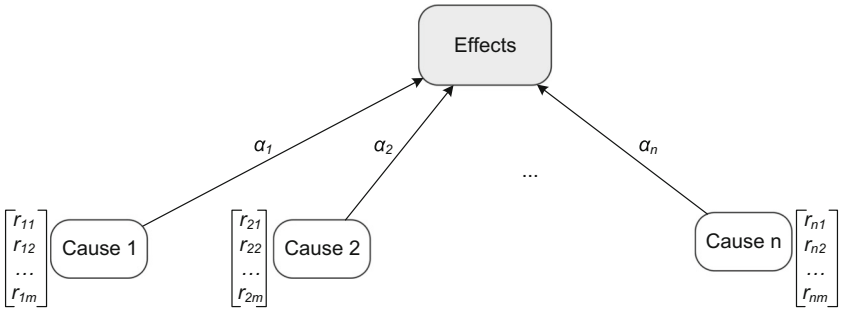

Fig. 2 The cause-effect structure and the causes' relation weights

has state of $j_1, j_2, ..., j_n$ respectively; and $\sigma(j_i - j, j)$ is calculated as:

$$\sigma(j_i - j, j) = \begin{cases} j_i - j, j_i - j \geq 0 \\ j, j_i - j < 0 \end{cases}$$

Besides, $I(C_i)$, the influence of C_i to the *effect*, can also be calculated by formula (2), which is obtained from [16]. In the formula, $P(E = m|C_i = k)$ is the conditional probability when Effect E has state m (highest) and cause C_i has state k. Element has the highest $I(C_i)$ value is the most influential element to the effects.

$$I(C_i) = \frac{\sum_{k=1}^{m-1} \frac{P(C_i=k)}{\sum_{j=1}^{m-1} P(C_i=j)} P(E=m|P(C_i=k)) - P(E=m|C_i=m)}{P(E=m)} \quad (2)$$

2.3 Incorporating FAIR's LEF assessment into Bayesian network

FAIR, as can be seen in Fig. 1, is a structural analysis framework, which consist of three pairs of cause-effect relations, including [cause: C, A; effect: TEF], [cause: Tcap, CS; effect: V], and [cause: TEF, V; effect: LEF]. By incorporating FAIR into BN, we can derive the numerical output for the state inputs, hence extend a number of functions, e.g. to rank the LEF more rapidly, or to assess using fuzzy inputs. There are no probabilistic data in FAIR that can be transformed directly to BN. However, when the experts construct the FAIR framework, they have embedded such data and their knowledge into the FAIR tables. As a result, these tables can be used as the evidence to obtain the probabilistic relations for the BN. Our approach is to obtain the required parameters from the FAIR tables to create a BN following the approach presented in Section 2.2. In detail, we need to acquire the *fuzzy judging vector* and *weight vector* from the FAIR tables.

The FAIR framework is proposed to simplify the complexity of the analysis; hence, many data and knowledge have been excluded from the FAIR tables. As a result, it is impossible to provide an exact numerical model reflecting the framework. Instead, we aim at using the fuzzy evaluation to reflect the FAIR results numerically, therefore giving more concrete picture of how threats' LEF can be compared.

As the structures of the three cause-effect pairs of FAIR are similar, we will present the procedure to incorporate their general structure. Following this procedure, all the three FAIR analysis can be incorporated to the BN.

For the general structure, assume that we have **effect** E and two **causes** C_1, C_2 each can have one of the five states [VL L M H VH]. The corresponding FAIR look-up table for E is [$e_{ij} \in$ {VL, L, M, H, VH}, $i = 1...5$, $j = 1...5$], entry e_{ij} is the result of input C_{1i} and C_{2j}. Our method to acquire the probabilistic relations consists of the steps illustrated in Fig. 3, details as follow:

Step 1 Calculating the *weight vector*: Noteworthy, the FAIR tables are formed with the assumption that the states of the two causes create direct impacts to the state of the effect. Therefore, if we transform the state data to numerical data, there should be a strong correlation between the cause and effect data in most of the cases. On the other hand, if the correlation test of a table shows not significant, that only means that this table was formed with a different model rather than with FAIR. In the simplest form, we can assume the relation is linear and translate the node state into number by defining VL=1; L=2; M=3; H=4; VH=5. We then have numerical data for the causes and effect, which we can use to run a regression to test the linear model between the causes and effect, $E = \alpha C_1 + \beta C_2 + \delta$ (α and β are the coefficients and δ is the error). The coefficients α, β are then standardized with $\alpha' = |\alpha|/(|\alpha| + |\beta|)$ and $\beta' = |\beta|/(|\alpha| + |\beta|)$. We choose $w = [\alpha' \ \beta']$ as the weight vector for the BN model. $\alpha' > \beta'$ indicates that cause C_1 has more significant impact to E than C_2, which means that a change in the C_1 value will fluctuate the E output more than the same change in C_2 input.

Step 2 Calculating the *fuzzy judging vector*: this vector judges and compares the impact of each state of the cause on the effect. For the FAIR framework, the low value of the input state will normally lead to the low value of the output state and vice versa. As can be seen from

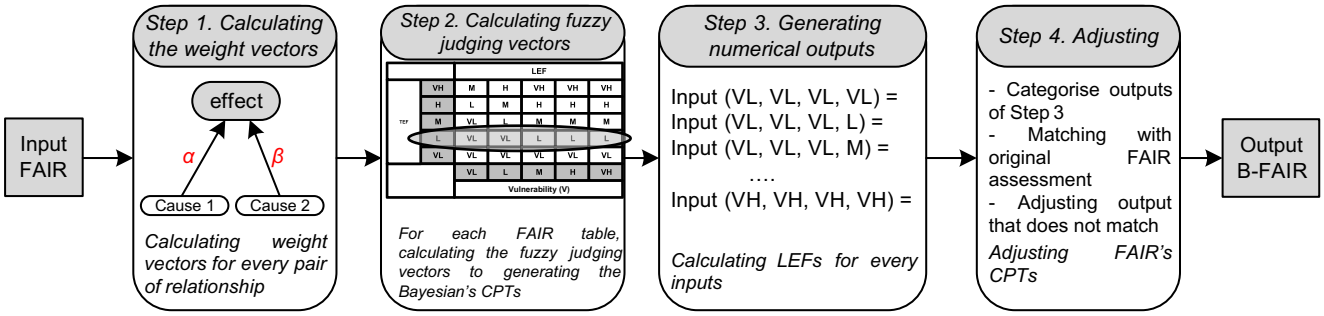

Fig. 3 Flow chart of generating B-FAIR: incorporating FAIR to Bayesian Network

Fig. 4, the FAIR table provides all the potential outputs that can be derived from a single state. For example, a TEF with L value can only lead to a LEF with VL or L value. These potential outputs can be used for acquiring the comparison of the impact of each state of the cause.

In order to sharpen the difference between the levels of the state, we convert further state e_{ij} to n_{ij} in which $n_{ij} = k^{e_{ij}}, k > 1$. So we have $n(VL) = k, n(L) = k_2, n(M) = k_3, n(H) = k_4$, and $n(VH) = k_5$. We also set the weights for the state of the cause *(VL, L, M, H, VH)* as $(\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5)$ to further differentiate the effect of the state from the other cause (see Fig. 2). The choice of k and the state weights will not affect to the correctness of the ranking orders that FAIR can verify. In detail, we assume that FAIR can always rank effects that have same input values in one factor. For example, in the fourth line of Fig. 3, the ranking order of LEF should be $LEF(TEF = L, V = VL) < LEF(TEF = L, V = L) < ... < LEF(TEF = L, V = VH)$ because the input value of factor TEF is always L, while the input value of V factor is in ascending order. In such cases, the choices of k and state weights will not affect to the order as long as $n_{21} < n_{22} < ... < n_{25}$. The larger the value of k, the deeper the numerical difference between evaluation output of the threats will be. For each cause, we derive its *fuzzy judging vector* $r = [r(VL), r(L), r(M), r(H), r(VH)]$ by calculating the individual effect value of each state s_i as: $r(s_i) = \frac{\sum_{j=1}^{5} \gamma_i n_{ij}}{\sum_{i=1}^{5}\sum_{j=1}^{5} \gamma_i n_{ij}}$.

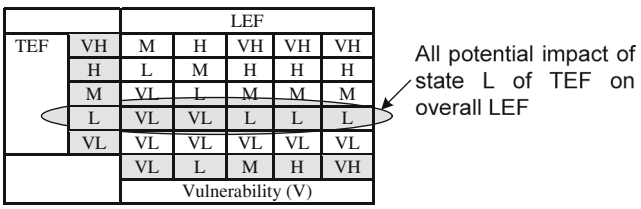

Fig. 4 Justifying the *fuzzy judging vector*

For each of the relations, after obtaining the *weight vector* and the *fuzzy judging vector*, we can generate the Bayesian CPT in each of the effect node following the formula in Section 2.2. Having the three CPTs from the three FAIR look-up tables is enough to form the overall Bayesian network for calculating the *LEF* output given the input states of the causes.

Step 3 Generating numerical output: The output of the Bayesian will be a vector of the probability of the state evaluations for the *LEF*, for example, $[p_1, p_2, p_3, p_4, p_5]$, in which p_1 is the probability that *LEF* has state *VL*, p_2 is the probability that *LEF* has state L, and so on. We use the grade vector $[1, 2, 4, 8, 16]$ to derive the final numerical result, in detail, the assessment for *LEF* is equal to $p_1 + 2 * p_2 + 4 * p_3 + 8 * p_4 + 16 * p_5$. This grade will later be used to compare and rank the threat, according to their *LEF*.

Step 4 Adjusting Bayesian model for FAIR consistency: Sometimes there may have some inconsistencies between the FAIR and Bayesian model due to the weak correlation of the values in the FAIR table. For example, with the same input state, FAIR output gives a "Low" state, but Bayesian does not give a low numerical output. In such cases, we provide fixed by adjusting the corresponding CPT entry of the Bayesian model based on the upper/lower bound grade according to the FAIR state. In detail, we group 25 FAIR outputs for LEF into five categories [VL L M H VH]. In each category, we will replace the FAIR output by the corresponding Bayesian grade (with the same input). We then obtain the value range for each category. If there is no intersection between the value ranges, the Bayesian model is fully consistent with the FAIR assessment. In case there are intersections, we will decrease the upper bound (for instance, decrease to the same value with the second highest upper bound in the same category) or increase the lower bound of the relevant categories accordingly to eliminate all the intersections. We then update all the CPT entries that related to the adjustments using the BN sensitivity analysis to ensure the adjustment does not affect the overall BN. After this stage, we can ensure the

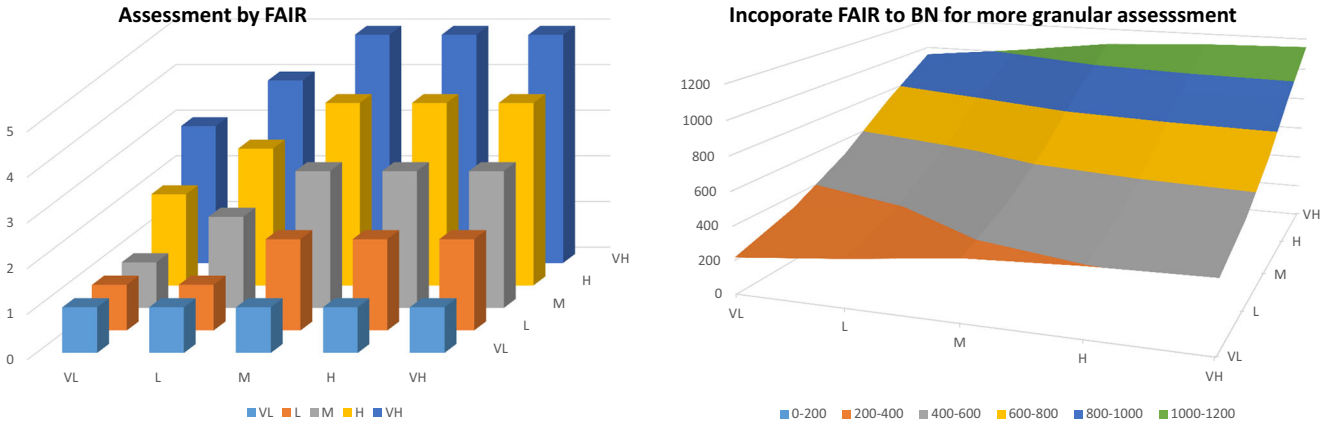

Fig. 5 Comparison between FAIR and FAIR incorporated into BN

consistent assessments with all the 25 inputs that FAIR can provide.

Figure 5 illustrates the comparison between FAIR and FAIR incorporated into BN (BN-FAIR) for the structure with causes [TEF V] and effect LEF. It can be seen that FAIR only evaluates threats with fixed and concrete inputs, while the BN-FAIR can estimate threat severity with flexible and consecutive inputs. Moreover, the BN calculations adjust the FAIR output in a way that threats in the same category can be compared together based on the numerical appraisal, yet still classify threats like FAIR does. The figure shows how sensitive the outputs are given changes in the inputs.

In the next section, we demonstrate how our approach extends FAIR by applying it to assess the potential smart grid threats for our Improving the Robustness of Urban Electricity Network (IRENE) research project [11].

3 Experiment results and discussions

3.1 Applying the method: context

The IRENE project is devoted to the topic of developing a collaborative framework to deal with the cyber threats to improve the urban grid resilience. To account for the threat landscape, the list of 102 information security threats outlined by NIST 800−30 [4] was refined into 38 potential threats. These threats can directly affect smart city operation, as described in [11]. In this paper, due to limited space, we only present how our method assesses the LEF of 14 threats from this 38-threat list, as can be seen in the second column in Table 1. The FAIR tables are taken from [12].

Figure 6 shows how the described BN-FAIR approach can be applied. The adopted approach named Threat Navigator aims to help stakeholders concentrate on threats with high LEF in a traceable and repeatable way, thus reducing the number of threat events that need to be further analysed. The logic of the approach is closely related to the logic of the state-of-the-art Intel's [17]. The approach is described in detail in [18]. In summary, a number of inputs (numbered 1 to 5 in the figure) are pre-processed to construct: a list of threats to be refined, Actor-to-Asset and Threat-to-Threat connections; and Threats-to-Mitigations links. Subsequently, the Threat Navigator employs pre-defined relations to remove threats less relevant to specific classes of attackers based on their Focuses and Capabilities. Next, the method looks for implemented mitigations. Finally, it employs the method outlined in this paper to calculate LEF for threats. Figure 6 provides a high-level view of the method and outputs provided by the BN-FAIR approach.

The procedure to obtain the input state for each of the factors is complex, so it will not be covered here. Readers are referred to [19] for more detail of our input state evaluation method. Let us assume that after the evaluation, the inputs for the 14 threats are given in the third column of Table 3. Among the input, threat 9 and 13 have fuzzy values. This is because for threat 9, the security experts were not able to assign whether "M" or "H" state for the "Tcap" factor. The chosen value indicates that we give a 40% belief for the "M" state and 60% for the "H" state. For threat 13, the experts were not be able to evaluate the "A" factor at all, so we choose the equal probability for each state. Although FAIR does not support assessments in these two cases, our method enables to benefit from the FAIR approach even with such input. The details of applying the method to the LEF element of FAIR taxonomy and propagating the data to obtain output values are described next.

Table 1 Numerical results of the Bayesian-FAIR to compare with FAIR

ID	Name	Input state	F	BF	R	M
1	Perimeter network scanning	[M, H, M, M]	H	889.5	7	C
2	Information gathering	[VH, H, M, H]	H	1016.9	4	C
3	Reconnaissance	[M, M, VL, L]	L	571.7	10	A
4	Craft phishing attacks	[H, H, VH, H]	H	1130.3	3	A
5	Spyware/Malware	[M, VH, H, VL]	VH	1147.1	1	A
6	Sniffers/Scanning	[M, H, H, M]	H	923.1	6	C
7	Insert subverted individuals	[M, H, H, VL]	H	939.9	5	CS
8	Exploit physical access	[L, M, L, H]	VL	342.9	13	A
9	Exploit unauthorized access	[H, M, 0.4M-0.6H, VH]	n/a	290.33	14	A
10	Exploit split tunneling	[L, H, H, M]	M	685.1	9	C
11	Exploit mobile systems	[VH, H, VH, H]	VH	1147.1	1	C
12	Exploit recently vulnerabilities	[H, M, H, VH]	M	809.7	8	A
13	Physical compromise	[VL, E, L, VL]	n/a	343	12	A
14	Hardware compromise	[M, L, H, M]	VL	397.4	11	A

F: FAIR approach; BF: Bayesian FAIR approach; R: Rank; M: Most Influential Factor; E: indicates the equal probability of 20%VL – 20%L – 20%M – 20%H – 20%VH (input of threat 13)

3.2 Results and discussions

Following the first step in Section 2.3, we obtain the *weight vectors* as follows $w(C, A) = [0.39\ 0.61]$; $w(Tcap, CS) = [0.5\ 0.5]$; $w(TEF, V) = [0.7\ 0.3]$. Note that the Pearson correlation tests for these three pairs returned strong statistical correlations (5% statistical significance), which confirmed that the linear regression model fitted to the FAIR table data. Choosing $k = 2$ and $(\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5) = (1, 2, 3, 4, 5)$ for *Step 2* we have the *fuzzy judging vectors* as follows: $r(C) = [0.42, 0.34, 0.18, 0.05, 0.01]$; $r(A) = [0.49, 0.34, 0.13, 0.03, 0.01]$; $r(Tcap) = [0.49, 0.3, 0.15, 0.05, 0.01]$; $r(CS) = [0.01, 0.05, 0.15, 0.3, 0.49]$; $r(TEF) = [0.62, 0.25, 0.09, 0.03, 0.01]$; $r(V) = [0.37, 0.3, 0.23, 0.08, 0.02]$. We classified and adjusted the BN's CPTs as described in *Step 3* and *4* to construct the final Bayes-FAIR model.

We then use the constructed Bayesian model to calculate the evaluation grade for the threats. The detailed results are given in Table 1. To see how the change in the belief of fuzzy inputs can change the overall assessment of a threat, we vary the belief for the input of the "A" state for threat 13, while the other three factors [C, Tcap, CS] are fixed to [VL, L, VL]. The changes are represented in Fig. 7, in which we calculate different evaluation grade when the input of "A" changes from [100%VL] to [20%VL 20%L 20%M 20%H 20%VH], [40%VL 15% L 15%M 15%H 15%VH], [60%VL 10%L 10%M 10%H 10%VH,, and [100%VH]. The lower bound, which is the lowest value of the calculated set, is 264.49, happened when "A" is at 100% "VL", while the upper bound is 531.3 when "A" is 100% "VH". The sensitivity of outputs given input changes can be seen through Fig. 7.

From Table 1, it can be seen that a Bayesian network, constructed according to the proposed method, generates assessment consistent with the FAIR framework. This is because the CPTs are derived from the FAIR look-up tables and can be adjusted for ensuring consistency. Moreover, our approach can differentiate further threats in the same category. For example, threat 6 and 7 are in the same "High"

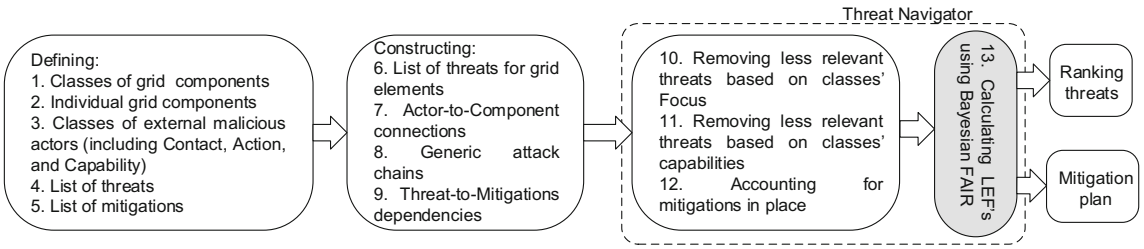

Fig. 6 Application of BN-FAIR approach

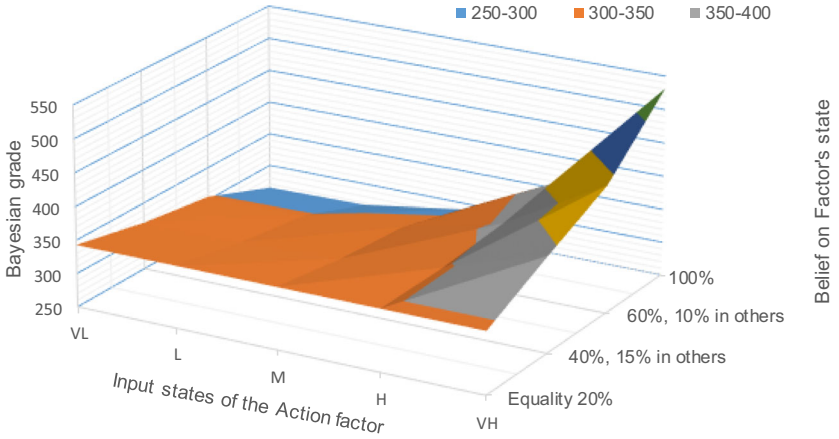

Fig. 7 Evaluation of LEF by Bayesian-FAIR with fuzzy state input of "Action" factor3.4 Discussion

category, according to FAIR, but having the grade of 923.1 and 939.9 respectively according to our approach. From the table, we can see that 6 and 7 have the same assessments for the three inputs [C, A, Tcap], the only difference is the evaluation of factor CS. Threat 7 has VL state compares to M of threat 6, so LEF of 7 should be higher than LEF of 6. This difference cannot be shown by FAIR as both of the threats are in the category, but it can be seen clearly from our Bayesian model.

In addition to providing a repeatable and traceable way to reach some conclusions, even in case of uncertainties, we supply a clear mechanism for integrating a threat threshold. Having the threat grades, we can simply define the cut out point to reduce the list of threats to consider. For example, a cut out point of 900 means threat is only considered when its grade higher or equal to 900, will reduce the list of threats to 2, 4, 5, 7, 6, 11.

Our model can also show the capability of giving output even with fuzzy input. For example, in case of threat 9 and 13, we can give the assessment grades of 290.33 and 343 respectively, while the FAIR model cannot provide the exact state. This capability will be helpful in the cases where there is a lack of expert opinions on assessing the threats, or experts have conflicted assessments of the threats. Another advantage is that our approach can point out the most influential factor for each of the threats to assess. These outputs then can be combined to show which factor should be improved to lower the threat impact. For example, considering the 14 threats in Table 1, we see that the "A" factor is the one that affect the most, with 8/14 of the threats. This suggests the system managers to implement some countermeasures to lower the "Action", for example, create some policies, which put higher punishment on the attackers that initiate such threats, to lower the motivation of attacking. Such countermeasures will lower significantly the impacts of eight threats in the list; hence, effectively improve the security system with the least efforts considering the 14 threats relevant for a particular smart grid configuration.

The outlined approach illustrates how to construct a continuum of LEF values within an individual class (e.g., High, Moderate, ...), as well as across several ranks. Noticeably, by doing so it provides opportunities, but not requirements, to rank threats across classes. This approach does not define that there must be a thin line between, for instance, High and Moderate classes. An analyst, equipped with the described solution, is expected to consider groups of threats differently and use the solution for its primary purpose – to numerically assess ranks within the same LEF category. Readers are referred to [19] for further details of a practical application of the Bayesian-FAIR to our IRENE project.

4 Conclusion

This paper presented a method to incorporate the FAIR structural analysis into the BN to obtain the LEF numerical threat assessment. The constructed BN-FAIR reflects and conserves the FAIR appraisal as all of the probabilistic relations are inferred from the FAIR tables. The numerical assessment is essential for the FAIR framework to further rank when there are too many threats in the same category, which will be more and more common in the future. Our method also extends the framework by allowing input fuzzy values when experts have conflicted evaluations, observe sensitivity of outputs given input changes, and identify the most influential factor to improve the mitigation effectiveness. We showed how the Bayes-FAIR can be applied in the IRENE research project and demonstrated its extended functions. We believe that this Bayes-FAIR can help the risk manager to formulate a more effective mitigation plan, which includes the most cost-effective security countermeasures to lower the threats' impacts. In

the future, we will focus on extending this method for the smart grid risk assessment.

References

1. Lo C-H, Ansari N (2012) The progressive smart grid system from both power and communications aspects. IEEE Commun Surv Tutorials 14:799–821
2. Knapp ED, Samani R (2013) Applied cyber security and the smart grid: Implementing security controls into the modern power infrastructure elsevier science
3. Wang W, Lu Z (2013) Cyber security in the Smart Grid: Survey and challenges. Comput Netw 57:1344–1371
4. Yan Y, Qian Y, Sharif H, Tipper D (2012) A survey on cyber security for smart grid communications. IEEE Commun Surv Tutorials 14:998–1010
5. Farahmand F, Navathe SB, Sharp GP, Enslow PH (2005) A management perspective on risk of security threats to information systems. Inf Technol Manag 6:203–225
6. Sun L, Srivastava RP, Mock TJ (2006) An information systems security risk assessment model under the Dempster-Shafer theory of belief functions. J Manag Inf Syst 22:109–142
7. Peltier TR (2005) Information security risk analysis. CRC press, Boca Raton
8. Shameli-Sendi A, Aghababaei-Barzegar R, Cheriet M (2016) Taxonomy of information security risk assessment (ISRA). Comput Secur 57:14–30
9. Jones J (2006) An introduction to factor analysis of information risk (fair). Norwich Journal of Information Assurance 2:67
10. McBride AJ, McGee AR (2012) Assessing smart grid security. Bell Labs Technical Journal 17:87–103
11. Jung O, Besser S, Ceccarelli A, Zoppi T, Vasenev A, Montoya Morales AL et al (2016) Towards a Collaborative Framework to Improve Urban Grid Resilience. In: presented at the IEEE International Energy Conference ENERGYCON 2016, Leuven
12. Stoneburner G, Goguen AY, Feringa A (2002) Sp 800-30 risk management guide for information technology systems
13. RMI (2007) FAIR Basic Risk Assessment Guide. Available at: http://www.riskmanagementinsight.com/media/docs/FAIR_brag.pdf
14. Aven T (2016) Risk assessment and risk management: Review of recent advances on their foundation. Eur J Oper Res 253:1–13
15. Zio E, Aven T (2011) Uncertainties in smart grids behavior and modeling: What are the risks and vulnerabilities? How to analyze them? Energy Policy 39:6308–6320
16. Dui H-y, Zhang L-L, Sun S-D, Si S-B (2010) The study of multi-objective decision method based on Bayesian network. In: 2010 IEEE 17Th international conference on industrial engineering and engineering management (IE&EM), pp 694–698
17. Rosenquist M (2009) Prioritizing information security risks with threat agent risk assessment, Intel Corporation White Paper
18. Vasenev A, Morales M, Ceccarelli A, Le A, Ionita D (2016) Threat navigator: grouping and ranking malicious external threats to current and future urban smart grids. In: presented at the 1st EAI International Conference on Smart Grid Inspired Future, SmartGIFT, in press
19. Vasenev A et al D2.2 — Societal impact of attacks and attack motivations, Improving the Robustness of Urban Electricity Networks IRENE, Available at: http://ireneproject.eu/wp-content/uploads/2016/06/IRENE-D2.2.pdf