Peng Liu
Sjouke Mauw
Ketil Stølen (Eds.)

LNCS 10744

# Graphical Models for Security

4th International Workshop, GraMSec 2017
Santa Barbara, CA, USA, August 21, 2017
Revised Selected Papers

Springer

# Lecture Notes in Computer Science 10744

*Commenced Publication in 1973*
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

More information about this series at http://www.springer.com/series/7410

Peng Liu · Sjouke Mauw
Ketil Stølen (Eds.)

# Graphical Models for Security

## 4th International Workshop, GraMSec 2017
Santa Barbara, CA, USA, August 21, 2017
Revised Selected Papers

Springer

*Editors*

Peng Liu
Pennsylvania State University
University Park, PA
USA

Ketil Stølen
SINTEF ICT Blindern
Oslo
Norway

Sjouke Mauw
University of Luxembourg
Esch-sur-Alzette
Luxembourg

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Computer Science
ISBN 978-3-319-74859-7 ISBN 978-3-319-74860-3 (eBook)
https://doi.org/10.1007/978-3-319-74860-3

Library of Congress Control Number: 2018930744

LNCS Sublibrary: SL4 – Security and Cryptology

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

# Contents

# CSIRA: A Method for Analysing the Risk of Cybersecurity Incidents

Aitor Couce-Vieira[1,3](✉), Siv Hilde Houmb[2], and David Ríos-Insua[3]

[1] Universidad Rey Juan Carlos, Madrid, Spain
am.couce@alumnos.urjc.es, aitor.couce@icmat.es
[2] Secure-NOK AS, Stavanger, Norway
sivhoumb@securenok.com
[3] Consejo Superior de Investigaciones Científicas,
Instituto de Ciencias Matemáticas, Madrid, Spain
david.rios@icmat.es

**Abstract.** Analysing risk is critical for dealing with cybersecurity incidents. However, there is no explicit method for analysing risk during cybersecurity incidents, since existing methods focus on identifying the risks that a system might face throughout its life. This paper presents a method for analysing the risk of cybersecurity incidents based on an incident risk analysis model, a method for eliciting likelihoods based on the oddness of events and a method for categorising the potential ramifications of cybersecurity incidents.

**Keywords:** Cybersecurity · Risk analysis · Incident risk analysis
Decision support

## 1 Introduction

Cybersecurity incidents happen in a context of uncertainty in which incident responders have to analyse the potential uncertainties around the incident and the potential consequences in the system and on the assets. The earlier signs of one of these events are, typically, suspicious anomalies that could also be caused by legit actions by the system or users. Here, the analysis focuses on identifying what could have caused the anomalous event, and what events might follow. For instance, a new connection within a network could be caused by a maintenance laptop or an unauthorised party accessing the network. Additionally, if a specific attack or problem has been identified, then the analysis focuses on identifying the consequences of the threat, how likely they are or how the potential countermeasures would change the risk. For example, analysing the presence of malware in an industrial controller would deal with aspects such as whether it is harmful to the controller or the current industrial process, whether it can spread to other devices or what the consequences of removing the malware or changing the device are.

P. Liu et al. (eds.): GraMSec 2017, LNCS 10744, pp. 57–74, 2018.
https://doi.org/10.1007/978-3-319-74860-3_4

Methods for cybersecurity risk analysis may be classified into three approaches: upstream, downstream and combined. Upstream methods, such as attack trees [1], fault trees or probabilistic attack graphs [2], identify the causes of the main incident. Downstream methods, such as FMECA[1] [3] or event trees [4], identify the consequences of the main incident. Combined methods, such as bow-ties [5,6] and risk matrices [7], cover both upstream and downstream analysis. A bow-tie combines an upstream tree for the causing events of the main incident and a downstream tree for its consequences. Risk matrices assign an ordinal value to the likelihood and to the severity of a risk, and then derive an ordinal risk rating from both values. Other relevant combined methods are CORAS [8] and FAIR [9]. Most of the existing methods, especially upstream and downstream ones, concentrate on risk description[2] [10] rather on risk evaluation.

Risk matrices are the most popular risk analysis method, but its limitations [7] are even more problematic when it comes to analysing incidents. First, combining the qualitative interpretations of likelihood in a chain of events would become meaningless, since they do not follow probability axioms. Second, analysing the impact over assets with them also present problems. On the other hand, risk matrices are very suggestive on what stakeholders should value as most frameworks using them provide a supporting table identifying some impact categories (e.g., people, property, reputation) and the corresponding severity level. In addition, they are also very suggestive on how should stakeholders evaluate the risk, since most frameworks provide an already coloured matrix to categorise risks.

This paper presents a method for analysing the risk of cybersecurity incidents, hereafter called CSIRA. The method combines a general model for incident risk analysis, a model for categorising the ramifications of cybersecurity incidents and a minimal method for eliciting likelihoods based on the oddness of events. These methods are introduced in Sect. 2. Section 3 introduces CSIRA, supported by an example of its application in Sect. 4. Finally, Sect. 5 briefly discusses our contributions and future work.

## 2 Base Models

### 2.1 GIRA: A General Model for Incident Risk Analysis

Figure 1 depicts a general model for incident risk analysis (GIRA), represented as an influence diagram. GIRA [11] combines risk information from upstream and downstream risk descriptions, as well as risk evaluation. As an influence diagram, GIRA provides a visualisation of the cause-effect relations of the risk, and the capability of processing quantitative and qualitative elicitations of it (this last one through a semi-quantitative procedure).

[1] Failure mode, effects and criticality analysis.

[2] In ISO terminology, risk description is named risk analysis whereas risk analysis is named risk assessment.

In an influence diagram, ovals represent events with uncertain states ('what could happen?'). Double-lined nodes represent events with deterministic/known states ('what would happen?'). Rectangles represent a set of alternative actions that decision-makers can take ('what we can do?'). Hexagons represent a set of preferences over the outcomes of a node ('how we value what could happen?'). Arcs represent conditional relations between nodes ('if this happens in the antecedent, then that happens in the consequent'). Stacked nodes represent that for certain node types, there could be several of them.

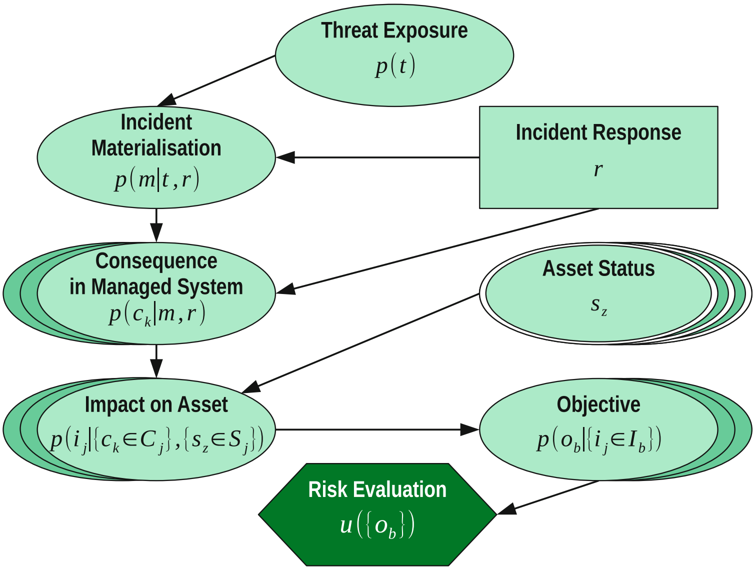


**Fig. 1.** GIRA depicted as an influence diagram.

The *threat exposure* node represents the likelihood that a threat is present in, or targeting, the system that the incident handlers are in charge of protection (MS, the managed system). Mathematically, it is represented by the probability distribution $p(t)$. The *incident response* node represents the alternative actions that the incident handlers could implement to avoid or mitigate the incident. The variable representing these actions is $r$. The *incident materialisation* node represents the likelihood that the threat materialises as an incident in the MS, taking into account the response of incident handlers. This is the first conditional node, $p(m|t,r)$, which means that the probability of incident materialisation depends on the threat presence and the response. The *consequences in the managed system* nodes represent the likelihood that an incident or its response cause further negative events in the MS. Its distribution is modelled as $p(c_k|m,r)$. There could be multiple nodes of this type, so we define the set of consequence nodes as $\{c_k\} = \{c_1, \ldots, c_K\}$, being $K$ the total number of consequences.

An asset is any element affected by the incident and valuable to the stakeholders. The *impact on asset* nodes provide the likelihood that a consequence in the MS leads to impacts over the assets of the MS or other systems, or over any other stakeholders' interests. This node takes into account the current *asset status*, which might enable or escalate the impacts of the incident. An asset status is represented as $s_z$ and the set of asset status nodes as $\{s_z\} = \{s_1, \ldots, s_Z\}$. An impact on asset node is represented as $p\Big(i_j|\{c_k : \exists\ c_k \rightarrow i_j\}, \{s_z : \exists\ s_z \rightarrow i_j\}\Big)$, being $\{c_k : \exists\ c_k \rightarrow i_j\}$ the set of consequence nodes parenting the $j$-th impact node[3] and, similarly, $\{s_z : \exists\ s_z \rightarrow i_j\}$ the asset status nodes parenting the $j$-th impact node. The set of impact on asset nodes is $\{i_j\} = \{i_1, \ldots, i_J\}$. The *objective* nodes synthesise impacts in a reduced number of objectives to facilitate stakeholders understanding and comparing the outcome of the incident. An objective node is represented as $p\Big(o_b|\{i_j : \exists\ i_j \rightarrow o_b\}\Big)$, being $\{i_j : \exists\ i_j \rightarrow o_b\}$ impact on assets nodes parenting the $b$-th objective node. The set of objective nodes is $\{o_b\} = \{o_1, \ldots, o_B\}$.

The combination of all the nodes, from threat exposure to objective nodes, represents *risk description*, which is modelled by the following equation:

$$\begin{aligned} & p\Big(\{o_b\}, \{i_j\}, \{s_z\}, \{c_k\}, m, r, t\Big) = \\ & \quad = p(o_1, \ldots, o_B, i_1, \ldots, i_J, s_1, \ldots, s_Z, c_1, \ldots, c_K, m, r, t) = \\ & \quad = \left[\prod_{b=1}^{B} p\Big(o_b|\{i_j : \exists\ i_j \rightarrow o_b\}\Big)\right] \left[\prod_{j=1}^{J} p\Big(i_j|\{c_k : \exists\ c_k \rightarrow i_j\}, \{s_z : \exists\ s_z \rightarrow i_j\}\Big)\right] \\ & \qquad \times \left[\prod_{k=1}^{K} p(c_k|m, r)\right] p(m|t, r)\ p(t). \end{aligned} \tag{1}$$

Finally, the *risk evaluation* node represents the stakeholders' evaluation of the risk scenarios caused by the incident. It can be modelled, following the multi-attribute utility theory paradigm [12], as $u\Big(\{o_b\}\Big) = u(o_1, \ldots, o_B)$. The actual risk evaluation is based on the expected utility when response $r$ is implemented,

$$\psi(r) = \int \ldots \int \quad u\Big(\{o_b\}\Big) p\Big(\{o_b\}, \{i_j\}, \{c_k\}, m, t\Big) \quad \mathrm{d}t\ \ \mathrm{d}m\ \ \mathrm{d}c_K \ldots \mathrm{d}o_1. \tag{2}$$

From this equation, we can obtain the maximum expected utility response, by calculating $r^* : \max \psi(r)$.

Another aspect to consider is the time frame of the risk analysis. Specifically, the *expiration time* ($e$) of GIRA is the estimated moment of the earliest relevant change in any of the elements that participate in the incident (e.g., threat, system, assets). The expiration time could also be a specific time frame set by the analyst. The analysts should refer likelihoods to such time frame.

[3] More properly, the set of consequence nodes for which there exist an arc (directed edge as a graph) directed to the impact node $i_j$.

## 2.2 Eliciting the Likelihood Based on the Oddness of the Event

The quality of risk analysis relies on how well it considers uncertainty [13]. This is achieved by using suitable and well-processed data, if available, or in the partial or complete support of expert knowledge [14] or other elicitation methods [15]. However, this information might not be available during the time frame of the incident, in which the analysts do not have access to data or experts.

Analysing the likelihood of events using a qualitative interpretation could be arbitrary, but a meaningful yet practical approach is basing this splitting on a qualitative interpretation of probability ranges: *certain* for $p(e) = 1$, *possible* for $p(e) = (t, 1)$, *rare* for $p(e) = (0, t)$ and *impossible* for $p(e) = 0$. Any event $x$ that clearly has a likelihood below the interpretative oddness threshold $t$ is defined as rare, whereas the events with a likelihood around or above $t$ are defined as possible. This simple method can be extended with several levels of *oddness*. Interpretatively, this means that rare would change to $p(e) = (t_2, t_1)$ and could be conceived as *rare (oddness 1)*, and that we could define a new *rarer than rare/rare (oddness 2)* event with $p(e) = (t_3, t_2)$. We can continue this process until a *are (oddness i)* event, which might be useful for comparing the likelihoods of different events, although it would become more and more difficult to interpret in absolute terms.

Additionally, we can establish a rule for the likelihood of a chains of $n$ events, based on the accumulated oddness, i.e.,

$$p(e_n|e_{n-1}|\dots|e_1) = (t_{l-1}, t_l) : l = \sum_{i}^{n} \mathrm{odd(e_i)},$$

being $\mathrm{odd}(e_i)$ the oddness of the event. Certain and possible events have an oddness of zero. Additionally, any chain with at least one impossible event is automatically impossible, and any chain with all of its events certain is automatically certain.

Following this rules we have that a chain o possible and certain events is possible, a chain with a rare event would be rare (one event with oddness 1), a chain with two rare events would be a rarer than rare event (two events with oddness (1), a chain with a rarer than rare event would be a rarer than rare event too (one event with oddness (2). For instance, in industrial cybersecurity, an analyst could interpret that the event of an attacker manipulating a controller is rare and that, given such a manipulation, the event of a controlled sabotage by the attacker is rare. Therefore, this chain of events would be elicited as 'rarer than rare event'.

Table 1 summarises these concepts. It also shows the numerical implementation in a Bayesian network like GIRA, which can take the qualitative likelihood as a numerical input to populate the probabilities of nodes and, vice versa, translate the overall probabilities calculated by the network into the qualitative interpretation again. These values are defined based on practical purposes. First, a probability range of 2 orders of magnitude, e.g. $(1 \times 10^{-2}, 1)$, allows us to model dozens of states. The differences among the magnitudes of the various probability ranges are established in a way so that a chain with a rare event will always

**Table 1.** Table with the probabilistic interpretation of qualitative likelihoods.

| Qualitative likelihood | Probabilistic interpretation | Numerical input to GIRA Bayesian network | Numerical output from GIRA Bayesian network |
|---|---|---|---|
| Certain | 1 | 1 | 1 |
| Possible | $(t_1, 1)$ | $(1 \times 10^{-2}, 1)$ | $(1 \times 10^{-10}, 1)$ |
| Rare (oddness 1) | $(t_2, t_1)$ | $(1 \times 10^{-12}, 1 \times 10^{-10})$ | $(1 \times 10^{-20}, 1 \times 10^{-10})$ |
| Rarer than rare (oddness 2) | $(t_3, t_2)$ | $(1 \times 10^{-22}, 1 \times 10^{-20})$ | $(1 \times 10^{-30}, 1 \times 10^{-20})$ |
| ... | ... | ... | ... |
| Impossible | 0 | 0 | 0 |

have a lower probability than a chain without it. In the case of GIRA, we have a chain of 5 nodes and, taking into account that we use probability ranges of 2 orders of magnitude, the difference between probability ranges must be, at least, 10. This way, by multiplying the probabilities of the chain of events, we will get as output the overall probabilities, with their different orders of magnitude.

### 2.3 Understanding Potential Ramifications of Cybersecurity Incidents

Multiple guidelines and taxonomies exist for identifying and categorising cybersecurity risks. We can distinguish two groups. One group at the technical level, the larger in the literature, deals with the categorisation of cyber attacks and their effects on digital systems. These guidelines might be useful for identifying elements related to threats, incidents, and system consequence. The other group deals with the impact that cybersecurity risks might have on assets, value or risk objectives. Examples of widely used methods are COBIT [16] or FAIR [9]. However, the majority of the categories for impacts and assets have a perspective that pivots on a business/organisational interpretation of assets and stakeholders. Although most risk management happens in organisational settings (e.g., business or public agencies), a more broad perspective is feasible when thinking about cybersecurity risk impacts, i.e., asset as something with value for somebody and stakeholder as somebody that might be affected by the incident.

A thorough categorisation model would require a combination of IT, OT, cyber-phisical and cyber-psychological risks, an analysis of their impact at microsocial and macrosocial level and an analysis of what new cyber risks would emerge in the future (e.g. what risks the pervasive use of virtual reality will bring and how they could become cybersecurity risks). There is no scientific or technical literature so comprehensive. However, a simplified model for quick elicitation may be established. Figure 2 depicts a graphical model for categorising the potential ramifications of cybersecurity incidents. In the context of GIRA, this model might be helpful for identifying the consequences and impacts nodes.

The starting point is the MS, in which the analysed cybersecurity incident happens. The primordial risks of cybersecurity incidents are those involving the processing, storage and transmission of digital data. For example, ransomware, denial of service or man-in-the-middle attacks. These events could happen in the MS or other digital systems managed by the organisation dealing with the incident or third parties.

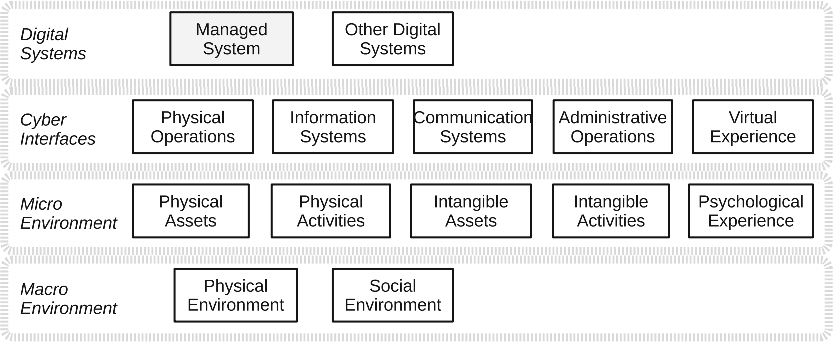


**Fig. 2.** Categories that classify the ramifications of cybersecurity risks

However, the importance of cyber risks resides, mostly, in the ramifications to other organisational or physical systems and assets that depend on, or can be affected by, the compromised digital systems. The most direct ramifications are the incidents grouped in the broad category of cyber interfaces. Physical operations refer to the interactions between physical reality and digital systems, such as input and output devices (e.g., keyboards, screens, printers, mouses, USB ports) or the actuators and sensors of industrial control systems. Examples of risks here involve unauthorised cyber-physical actions like the ones executed by Stuxnet [17] (manipulation of nuclear plant centrifuge speeds) or the malicious hijacking of laptop cameras. Information systems refer to the actual information contained in the digital systems (e.g., documents, pictures). An example risk in this case is the stealing of secret documents. Communication systems refer to the actual communication facilitated by the digital systems (e.g., chats, video conferences). Examples of risks here are the interference with a video conference or even the dissemination of false information through vulnerabilities in social networks (e.g., Twitter bots). Administrative operations refer to the affairs conducted with the digital systems (e.g., invoicing or buying online). An example risk in this area is the hijacking of an e-banking account. The virtual experience refers to the human experience in the reality created by the digital system (e.g., user experience in an application, human interaction in a social network). Examples of this type of risk are the exposure of personal information or sensitive images in social networks.

The indirect ramifications are categorised in a micro- and a macro-environment that refer to non-digital and non-cyber consequences. The micro-environment refers to risks at the particular or organisational level, as well as risks with organisations and people with a relatively direct relationship (e.g., customers and suppliers for a business, family and friends for a person). The first type of risks are in physical assets (e.g., machinery, personnel) and activities (e.g., manufacturing and transporting items). An example risk could be the sabotage by Stuxnet of the facility centrifuges (asset) and the enrichment of uranium (activity). Intangible assets refer to any characteristic or thing without physical presence. Example risks are the loss of secrets, reputation, compliance or money caused by a cyber attack. The psychological aspect refers to how cyber risks affect the human experience. Examples of these risks are the psychological problems derived from cyber-bulling or the exposure or personal data to the public. The macro-environment refers to the consequences at a social or ecosystem level. For instance, the political impact on Iran of Stuxnet, or the environmental and economic impact in the case a cyber attack facilitates an accident with contaminants or dangerous materials in an industrial facility.

## 3 CSIRA: Cybersecurity Incident Risk Analysis

Now we introduce the *cybersecurity incident risk analysis model (CSIRA)*, which aims at providing a paradigm practicable as a quick risk analysis method during cybersecurity incidents. CSIRA combines GIRA, the oddness method for likelihood elicitation, the graphical model for brainstorming cybersecurity incident ramifications and a simplified method for risk evaluation based on comparing the outcomes of different incident responses.

First, CSIRA uses GIRA (Sect. 2.1) as the risk analysis model, so that a high-level but comprehensive method is applied to the cybersecurity incident assessment. As argued previously, risk matrices oversimplify many risks components and other methods are more focused on the technical side (e.g. bow-ties). It is feasible to combine the use of a more detailed technical model for the cyber attack (e.g., attack tree) and the consequences (e.g., fault tree) with the use of GIRA for the impact and objective analysis.

Second, CSIRA uses a simplified interpretation of likelihood (Sect. 2.2), so that the elicitation is quick but at least implementable numerically. The qualitative scale of risk matrices cannot be applied to a chain of events nor be interpreted easily as a probability range. We also assume that a quantitative or semi-quantitative elicitation is not feasible in real-time. If so, then it would also be feasible to directly use GIRA, with quantitative data or expert elicitation.

Third, CSIRA uses a simplified model for eliciting the ramifications of cybersecurity incidents (Sect. 2.3), so that all feasible types of incidents are thought about. This intends to facilitate brainstorming, based the contextual knowledge of the user undertaking the analysis. We think that this approach is more feasible and useful in real time than presenting a general catalogue of impacts.

Fourth, GIRA would need the elicitation of the preferences and risk attitudes of the stakeholders, following the standard process in influence diagram building.

However, this would require time and support from experts. For CSIRA, we establish a faster alternative method, desicribed in Sect. 4.4: Once the users build the risk description part, they could obtain the total likelihoods of the risk problem. From the decision-making perspective, the only comparison they have to make is how the responses to the incident, and inaction, affect risk objectives.

CSIRA does not contain any knowledge base or any process to build one. For that to be useful, it would be necessary with very tailored information adapted to the specific systems, assets and stakeholders of the organisation. Indeed, rather than the potential incorporation of cybersecurity knowledge, we would recommend the use of a collection of cybersecurity standards. The most relevant one in this case is the NIST Cybersecurity Framework [18], which provides (1) the most comprehensive structuring of the aspects that should be taken into account in cybersecurity management and (2) specific chapters that deal with these topics in other relevant collections of standards (e.g., NIST, ISO, COBIT). Nor do we provide any automatic reasoning mechanism besides the Bayesian calculation of likelihoods. Although automation would reduce human task load, it would also take decision-making from the users. Indeed, the intention is the opposite: providing a risk analysis model that explicitly relies, as much as possible, on human interpretation and decision-making.

## 4 An Example Cybersecurity Risk Analysis

This section introduces the steps for using CSIRA, supported by an example. Our intention is not to undertake a realistic risk analysis but to provide an example to show CSIRA. First, we cover risk description, which consists in three steps. The first step, in Sect. 4.1, is risk identification using the graphical model presented in Sect. 2.3 for identifying cybersecurity incident ramifications. The second step is risk elicitation (Sect. 4.2), using GIRA as the base risk model (presented in Sect. 2.1) with the elicitation method presented in Sect. 2.2 to generate the likelihoods of different events. The final step of the risk description is risk calculation, using also the mentioned elicitation method. The outcome of risk description are the relevant risk scenarios for decision-making: the potential results of the different incident responses regarding their relevant risk objectives. The risk analysis finalises with the risk evaluation of Sect. 4.4.

The example case is applied to the industrial control systems (ICS) of an oil and gas drilling rig, as this facility is a paradigmatic case of the physical and organisational ramifications that a cybersecurity incident could have. The incident would be the presence of a wiper malware in the system in charge of drilling the well. This kind of malware is capable of erasing data in the operating system (OS) boot records or critical files. Interestingly [19] some of the most notorious wiper cyber attacks, like Shamoon and BlackEnergy, targeted the oil and gas industry. The human-machine interfaces (HMI) of industrial systems are typically installed on top of popular OS like Windows. Therefore, a disruption in the HMI caused by a wiper might affect, to some extent, the industrial operation that the HMI helps to control. This involves that incident handlers should think about the ramifications of the incident on industrial operations and assets.

### 4.1 Risk Description: Identification

Figure 3 depicts the consequences and impacts of cybersecurity incidents, applying the method of Sect. 2.3 to our scenario of a wiper in a drilling rig. The managed system is the drilling ICS. The initial incident is the presence of the threat, i.e., the presence of the wiper malware in the ICS. The exposure to this threat could lead to the main incident, which is the execution of the wiper in the PC hosting the HMI software. The square represents the potential response of the incident handler. Given that a wiper could be a sophisticated tool, a full fresh re-installation of the HMI PCs would be a prudent response.

In case the wiper is successfully running in an HMI PC, the next consequence could be the disruption of the OS of the HMI PC. In addition, the incident response has also a consequence: a fresh installation of the HMI PCs would need to put the ICS under maintenance for 24 h. The next step is to identify the ramifications that the disruption could have beyond the ICS. The first one is the disruption in the human-machine interface, i.e., the disruption of the interaction between operator, ICS and industrial operation. This could lead to a

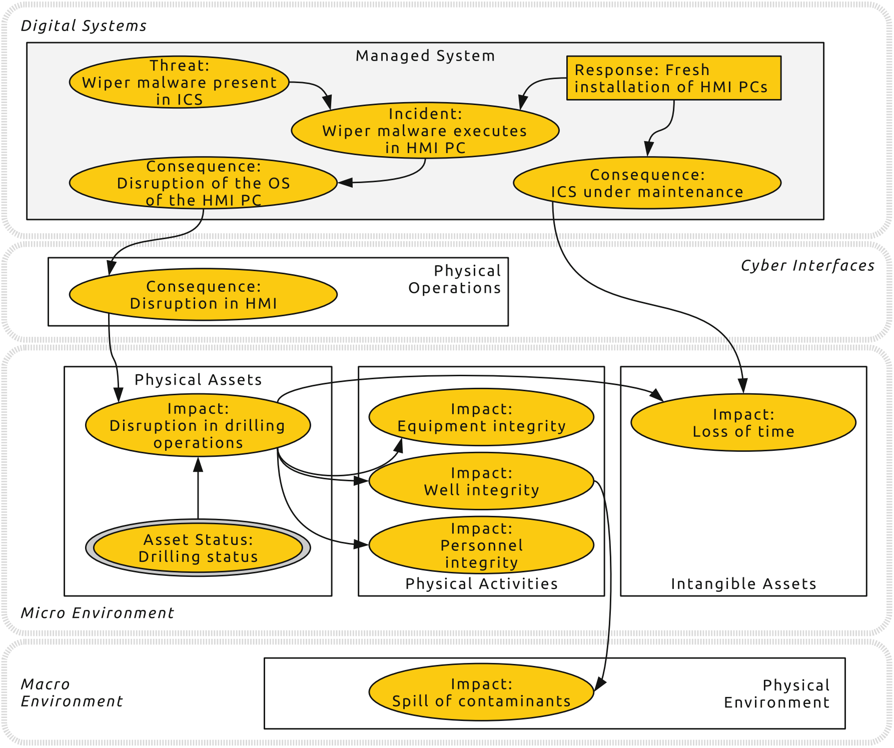

Digital Systems
Managed System
Threat: Wiper malware present in ICS
Response: Fresh installation of HMI PCs
Incident: Wiper malware executes in HMI PC
Consequence: Disruption of the OS of the HMI PC
Consequence: ICS under maintenance
Cyber Interfaces
Physical Operations
Consequence: Disruption in HMI
Physical Assets
Impact: Disruption in drilling operations
Asset Status: Drilling status
Impact: Equipment integrity
Impact: Well integrity
Impact: Personnel integrity
Physical Activities
Impact: Loss of time
Intangible Assets
Micro Environment
Macro Environment
Impact: Spill of contaminants
Physical Environment

**Fig. 3.** Graphical representation of potential risks of a wiper in a drilling rig. Rounded nodes represent uncertain events. Rectangles represent incident handler decisions. Double-rounded circles represent known states.

disruption of the drilling operations, which in turn might lead to incidents with equipment, the oil well or personnel. In addition, an incident involving the well integrity might lead to a spill involving hydrocarbons or other contaminants into the rig floor or the sea. An additional consequence, very relevant in oil platforms, is the loss of time, which can be caused by both the disruption in the drilling operations and the maintenance of the ICS (in the case of re-installing the HMI OS). However, one important element affects the disruption of the drilling operations: whether the platform is drilling or performing other activity.

### 4.2 Risk Description: Elicitation

Figure 4 illustrates the influence diagram of our example, using the likelihood elicitation of Sect. 2.2, and derived from the risks identified in Sect. 4.1.

The uppermost node is the *threat exposure*. It represents the uncertainty about the presence of the wiper. In this case, the analysts considered that the presence is possible (represented as P in the graph). Its complementary state (no presence of wiper) is also possible. Additionally, the *incident response* node represents the actions that the incident handler can take. In our case, the re-installation of the HMI OS with a fresh and updated version or the option of leaving the system as is.

The *incident materialisation* node represents the main incident: the execution of the wiper in the HMI PC. It has two uncertain states: whether the wiper runs in the PC or not. However, these events are conditioned by two factors. First, whether the wiper presence is a false alarm (threat exposure node). Second, whether the incident handlers re-install the HMI PCs. This is reflected in the likelihood assigned. If the wiper is present and the incident handlers leave the system as is, then it is possible that the wiper would run in the HMI PC. Otherwise, the wiper would not run (in the graph, 0 represents impossible and 1 represents certain).

There are two *consequence in the managed system* nodes. The first one represents the event of the wiper actually disrupting the OS of the HMI. In case the wiper is running in the HMI PC, then the likelihood of the HMI disruption is rare (as established earlier, rare (oddness 1), represented in the graph as R1) and the likelihood of its opposite is, thus, possible. In case the wiper is not running, then the certain event is the correct status. The second consequence node represents the event of putting the system under maintenance caused by the re-installation of the HMI PCs.

There are several *impact on asset* nodes. They represent most of the incident ramifications outside the managed system we identified in the previous section, except the disruption of drilling operations. The reason is that such disruption acts as an 'intermediate' risk, i.e., its risks are reflected on other assets, like the integrity of the different assets, the loss of time or the spill of contaminants. These nodes are preceded by the asset status node informing whether the platform is drilling. In addition, the impact nodes should summarise the likelihood of the chain of events that do not happen in the MS but may lead to those impacts. This means that given a consequence in the MS and the status of some asset,

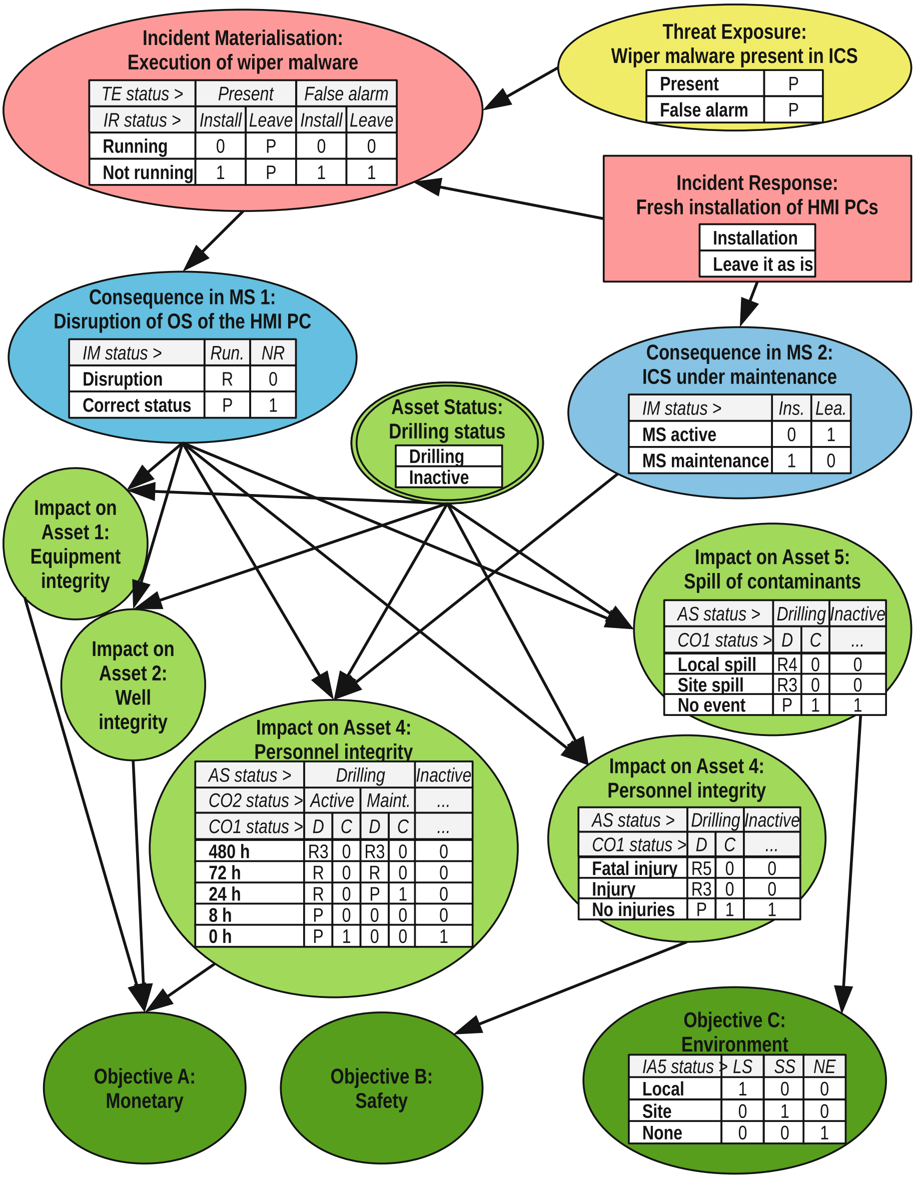


**Fig. 4.** Influence diagram representing the risk analysis for the wiper incident in a drilling control system. When it comes to the likelihoods, a sure event is represented with 1, an impossible event with 0, a possible event with P, a rare event with R, a rarer than rare event with R2, and so on.

they should reflect the likelihood of the different impact levels attainable. For instance, in case the impact 5 ‘spill of contaminants’ we have that, given that the asset status is drilling and that the HMI PC has been disrupted, the likelihood of a local spill is rare (oddness 4), the likelihood of a site spill is rare (oddness 3) and the likelihood of the no spill event is possible.

It is necessary to analyse the chain of events to determine whether one event is clearly rarer than other, as in Sect. 2.2. For instance, the event of a fatal personnel injury is established as clearly rarer than a non-fatal injury and than a local spill. Then, we establish that the event of a local spill is clearly rarer than a site spill. Following this procedure, we assign the different oddness to different events.

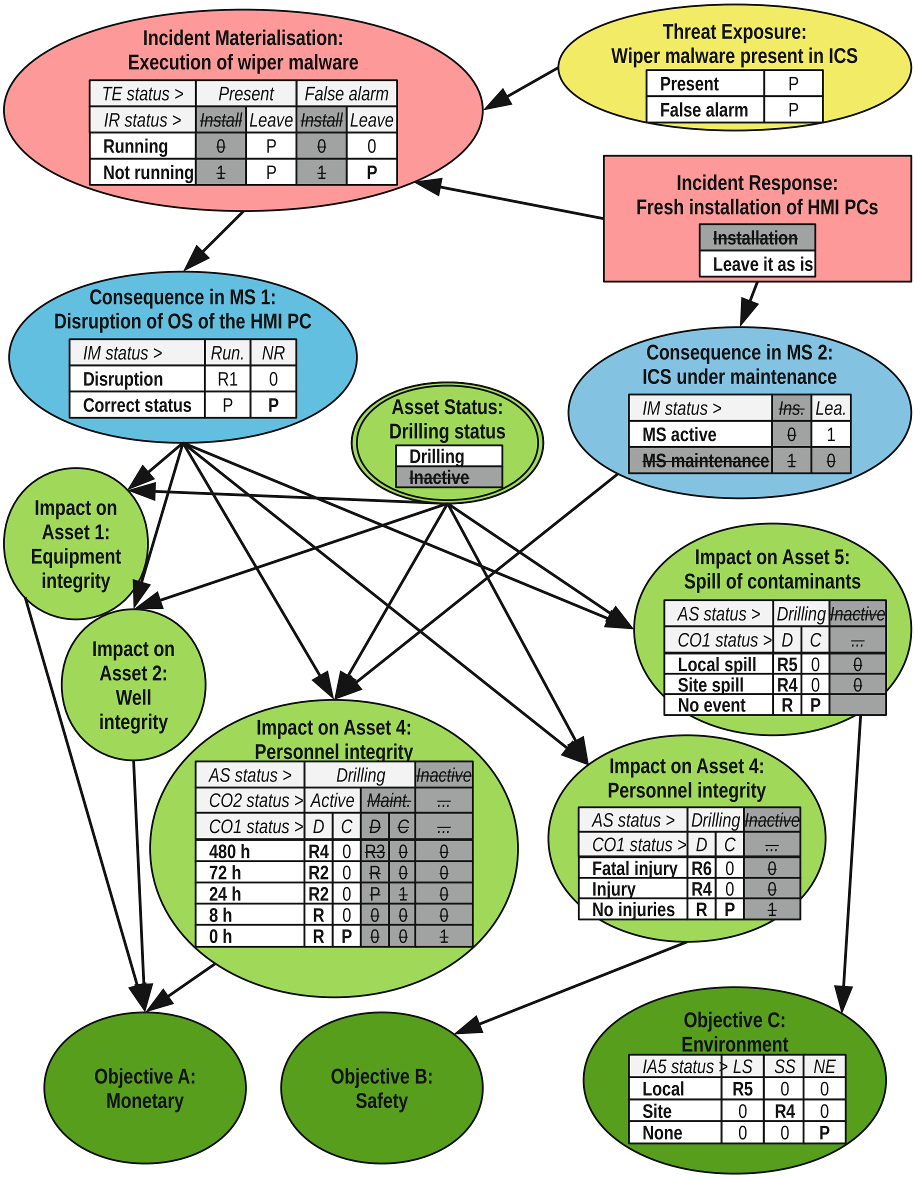
A final aspect to take into account is the expiration time of this risk analysis. Most of the events described have no clear time boundary. However, one of the nodes of our example stands out as the compass of timely risk response: the asset status node. First, all of the relevant impacts happen when the platform is drilling. Second, the incident handlers are able to know whether the platform is drilling or not and when this status would change. For instance, drilling might be scheduled for turns lasting several hours in the upcoming weeks. As an example, the expiration time for the analysis could be 8 h.

### 4.3 Risk Description: Calculation

Following the procedure for likelihood calculation in Sect. 2.2, we can calculate the final conditional probabilities of the different nodes of the influence diagram. Figure 5 displays the calculation for the case in which the incident response ‘leave the MS as is’ is selected and taking into account that the current asset status is ‘drilling’.

The logic of the influence diagram allows us to disregard infeasible and impossible events. For instance, the stricken out text in grey cells highlights infeasible events (e.g., in the consequence 2 node, it is infeasible any event that is conditioned by the incident response event of ‘installation’) or impossible events (once again, in the consequence 2 node, the event of ‘maintenance’ is impossible, given that the incident response event is ‘leave the system as is’). This kind of reasoning propagates through the diagram.

Additionally, the oddness method of likelihood propagation allows us to replicate conditional probability. For instance, in the incident materialisation node, the marginal likelihood of the event ‘wiper not running’, given the events ‘false alarm’ in the threat exposure node and ‘leave it as is’ in the incident response, is certain. However, its conditional probability is possible, since its materialisation is a chain of a possible event (‘false alarm’) and a certain event (‘wiper not running, given the false alarm and the leaving of the system as is’). This procedure propagates through the diagram. Additionally, when an event can happen through multiple event chains, then the likelihood of the likeliest one is selected. For example, in the impact on asset 5, the event ‘no spill event’ is rare if it comes from the chain with the consequence 1 event ‘disruption’, and it is possible if it comes from the chain with the consequence 1 event ‘correct status’. Since the event is, overall, at least possible, this is the likelihood passed to the child event ‘none’ in the objective C node.



| TE status > | Present | | False alarm | |
|---|---|---|---|---|
| IR status > | ~~Install~~ | Leave | ~~Install~~ | Leave |
| Running | ~~0~~ | P | ~~0~~ | 0 |
| Not running | ~~1~~ | P | ~~1~~ | **P** |

| | |
|---|---|
| Present | P |
| False alarm | P |

| |
|---|
| ~~Installation~~ |
| Leave it as is |

| IM status > | Run. | NR |
|---|---|---|
| Disruption | R1 | 0 |
| Correct status | P | **P** |

| |
|---|
| Drilling |
| ~~Inactive~~ |

| IM status > | ~~Ins.~~ | Lea. |
|---|---|---|
| MS active | ~~0~~ | 1 |
| ~~MS maintenance~~ | ~~1~~ | ~~0~~ |

| AS status > | Drilling | | ~~Inactive~~ |
|---|---|---|---|
| CO1 status > | D | C | ~~...~~ |
| Local spill | **R5** | 0 | ~~0~~ |
| Site spill | **R4** | 0 | ~~0~~ |
| No event | **R** | **P** | |

| AS status > | Drilling | | | | ~~Inactive~~ |
|---|---|---|---|---|---|
| CO2 status > | Active | | ~~Maint.~~ | | ~~...~~ |
| CO1 status > | D | C | ~~D~~ | ~~C~~ | ~~...~~ |
| 480 h | **R4** | 0 | ~~R3~~ | ~~0~~ | ~~0~~ |
| 72 h | **R2** | 0 | ~~R~~ | ~~0~~ | ~~0~~ |
| 24 h | **R2** | 0 | ~~P~~ | ~~1~~ | ~~0~~ |
| 8 h | **R** | 0 | ~~0~~ | ~~0~~ | ~~0~~ |
| 0 h | **R** | **P** | ~~0~~ | ~~0~~ | ~~1~~ |

| AS status > | Drilling | | ~~Inactive~~ |
|---|---|---|---|
| CO1 status > | D | C | ~~...~~ |
| Fatal injury | **R6** | 0 | ~~0~~ |
| Injury | **R4** | 0 | ~~0~~ |
| No injuries | **R** | **P** | ~~1~~ |

| IA5 status > | LS | SS | NE |
|---|---|---|---|
| Local | **R5** | 0 | 0 |
| Site | 0 | **R4** | 0 |
| None | 0 | 0 | **P** |

**Fig. 5.** Influence diagram representing the total conditional likelihoods for the risk analysis problem. Grey cells with the text stricken out represent infeasible or impossible events. Likelihoods in bold highlight that the conditional likelihood differs from the marginal one in Fig. 4.

### 4.4 Risk Evaluation

From an evaluative point of view risks and, specifically, impacts over value are incommensurable, i.e., they cannot, or ought not, be objectively evaluated in a single severity scale [20]. Therefore, a single scale, like the severity level of risk matrices, leads to a high level of incommensurability. On the other hand, it is recommendable to limit the number of elements to compare to facilitate decision-making. Multiple methods exist for evaluating risk, for instance, if the analyst has time and access to subject-matter experts, it is recommendable to use a method for preference and risk attitude elicitation, e.g. multi-attribute utility theory [21]. The rationality axioms make sense for generating a transparent and

| | Asset Status: | Rig is drilling | |
|---|---|---|---|
| | Response: | Leaving the MS as is | |
| Likelihood | Objective A: Monetary | Objective B: Safety | Objective C: Environment |
| Certain | - | - | - |
| Possible | € 0 | No injuries | No spill |
| Rare | € 80,000 | - | - |
| Rarer than rare | € 240,000<br>€ 720,000 | - | - |
| Oddness 3 or higher | € 2,000,000 [R3]<br>€ 4,800,000 [R4]<br>€ 5.000.000 [R5]<br>€ 6,800,000 [R7]<br>€ 9,800,000 [R9]<br>€ 11,800,000 [R12] | Injuries [R4]<br>Fatal injury [R6] | Site spill [R4]<br>Local spill [R5] |

| | Asset Status: | Rig is drilling | |
|---|---|---|---|
| | Response: | Fresh installation of HMI PCs | |
| Likelihood | Objective A: Monetary | Objective B: Safety | Objective C: Environment |
| Certain | € 240,000 | No injuries | No spill |
| Possible | - | - | - |
| Rare | - | - | - |
| Rarer than rare | - | - | - |
| Oddness 3 or higher | - | - | - |

**Fig. 6.** Tables representing the likelihood of different risk objectives when the incident handlers leave the wiper in the MS (upper table) and when they decide to do a fresh installation of the affected computers (lower table). Events with an oddness of 3 or higher contain their specific likelihood with squared brackets.

logical evaluation of the risk scenarios. Utility functions are flexible enough to represent multiple types of preference and risk attitudes and they offer strong analytical and mathematical properties. In addition, it is possible to avoid re-eliciting preferences as long as there are no changes in preferences.

The outcome of the risk description part is a set of scenarios representing how risk objectives could be affected by an incident, given the incident response. As depicted in Fig. 4, we created three objective nodes: monetary, safety and environment. The monetary node synthesises the cost that an incident in the assets might cause. On the other hand, the safety and environment nodes are practically direct translations of their precedent impact on asset nodes, as they have only one parent node.

As a decision problem, risk analysis is undertaken with the purpose of clarifying what are the best options to counter a risky situation. In our case, this involves that the main components to be evaluated are the potential responses of the incident handlers regarding risk objectives.

Tables in Fig. 6 display the relevant information that CSIRA presents to the stakeholders so that they are able to compare what different events regarding risk objectives, and their likelihood, might happen if they implement a response. In this case, the alternatives are either assuming a cost €240,000, caused by the lost time of maintaining the MS or face the rare event of losing €80,000, or the rarer than rare events of losing €240,000 or €720,000. If they disregard the even more rare events (oddness 3 or greater), then it seems a simple comparison between a certain lost of €240,000 and a loss three times greater but many more times less likely. However, should the stakeholders take into account the most rare events, then the comparison would become less clear.

## 5 Discussion

We have presented CSIRA, a model for building a high-level cybersecurity incident risk analysis. CSIRA is based on an influence diagram that provides a more comprehensive risk analysis than risk matrices. Realising the fact that risk quantification is practically infeasible in real time, we have implemented an alternative qualitative method that is at least implementable in an influence diagram to follow the basic logic of probability. We have put a special emphasis on what stakeholders value (impact nodes), how to synthesize these impacts over value (objective nodes) and how do stakeholders evaluate potential responses with respect to these risk objectives (risk evaluation). These axiological aspects require, rather than plain business impact scales, decision analysis modelling, so that value aspects are better formalised [22].

We present our method as an alternative to risk matrices rather than to more technical methods like attack or failure trees. Namely for two reasons, matrices use a single severity scale to merge their different categories of impact, in contrast to our approach or a more granular identification of impacts and their synthesis in a reduced number of risk objectives. Additionally, our likelihood elicitation method is as simple as the risk matrices (and it shares its limitations) but is designed to follow probability axioms, so that it could be applied to chains of events.

Upcoming work will focus on the implementation of CSIRA. Provide a more detailed specification of GIRA and the likelihood elicitation method. The main aspect is its software implementation. The R environment offers an ideal platform for elaborating a framework for the generation of CSIRA risk analysis case studies. Alternative, a Python implementation would facilitate the creation of an small application to undertake a CSIRA analysis. Besides the implementation of the influence diagram, that requires statistical and graph visualisation packages, it is also important to define a semantic model of CSIRA that captures the input from the users. Additionally, the elicitation method presented here would require a set of functions that transforms the user input (e.g., possible, oddness-1 rare event) into the marginal probabilities of the Bayesian nodes, and a set of functions that transforms the calculated probabilities into the 'oddness' language again. Future work after the implementation shall focus on test-based improvements of CSIRA and the construction of guidelines for its use.

**Acknowledgements.** The authors are grateful to the support of the MINECO MTM2014-56949-C3-1-R project, the AXA-ICMAT Chair in Adversarial Risk Analysis, the Regional Forskingsfond Vestlandet project 245291 Cybersecurity Incident Response Framework, and the COST IS1304 Action on Expert Judgement.

## References

1. Schneier, B.: Attack trees. Dr. Dobb's J. **24**(12), 21–29 (1999)
2. Singhal, A., Ximming, O.: Security Risk Analysis of Enterprise Networks Using Probabilistic Attack Graphs. National Institute of Standards and Technology, Gaithersburg (2011). https://doi.org/10.6028/nist.ir.7788
3. Department of Defense: MIL-STD-1629A, Procedures for Performing a Failure Mode, Effect and Criticality Analysis. Department of Defense, Washington DC, USA (1980)
4. Clemens, P.L., Simmons, R.J.: System Safety and Risk Management: A Guide for Engineering Educators. National Institute for Occupational Safety and Health, Cincinnati (1998)
5. International Association of Drilling Contractors: Health, Safety and Environment Case Guidelines for Mobile Offshore Drilling Units, Issue 3.6. International Association of Drilling Contractors, Houston, TX, USA (2015)
6. International Organisation for Standardization: ISO 17776:2000, Petroleum and Natural Gas Industries – Offshore Production Installations – Guidelines on Tools and Techniques for Hazard Identification and Risk Assessment. International Organisation for Standardization, Geneva, Switzerland (2000)
7. Cox, L.A.: What's wrong with risk matrices? Risk Anal. **28**(2), 497–512 (2008). https://doi.org/10.1111/j.1539-6924.2008.01030.x
8. Lund, M.S., Solhaug, B., Stølen, K.: Model-Driven Risk Analysis: The CORAS Approach. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-12323-8
9. The Open Group: Risk Taxonomy. The Open Group, Reading, UK (2009)

10. Cherdantseva, Y., Burnap, P., Blyth, A., Eden, P., Jones, K., Soulsby, H., Stoddart, K.: A review of cyber security risk assessment methods for SCADA systems. Comput. Secur. **56**, 1–27 (2016). https://doi.org/10.1016/j.cose.2015.09.009
11. Couce-Vieira, A., Insua, D.R., Houmb, S.H.: GIRA: a general model for incident risk analysis. J. Risk Res. (2017). Advance online publication https://doi.org/10.1080/13669877.2017.1372509
12. Keeney, R.L., Raiffa, H.: Decisions with Multiple Objectives. Cambridge University Press, Cambridge (1993). https://doi.org/10.1017/CBO9781139174084
13. European Food Safety Authority: Guidance on Uncertainty in EFSA Scientific Assessment. European Food Safety Authority, Parma, Italy (2016)
14. European Food Safety Authority: Guidance on Expert Knowledge Elicitation in Food and Feed Safety Risk Assessment. European Food Safety Authority, Parma, Italy (2014). https://doi.org/10.2903/j.efsa.2014.3734
15. Renooij, S.: Probability elicitation for belief networks: issues to consider. Knowl. Eng. Rev. **16**(3), 255–269 (2001). https://doi.org/10.1017/s0269888901000145
16. ISACA: COBIT 5: A Business Framework for the Governance and Management of Enterprise IT. ISACA, Rolling Meadows, IL, USA (2012)
17. Langner, R.: Stuxnet: dissecting a cyberwarfare weapon. IEEE Secur. Priv. **9**(3), 49–51 (2011). https://doi.org/10.1109/msp.2011.67
18. National Institute of Standards and Technology. Framework for Improving Critical Infrastructure Cybersecurity (2014)
19. Industrial Control Systems Cyber Emergency Response Team. Destructive Malware. National Cybersecurity and Communications Integration Center (US) (2014)
20. Espinoza, N.: Incommensurability: the failure to compare risks. In: The Ethics of Technological Risk, pp. 128–143. Earthscan, London (UK) (2009)
21. Reichert, P., Langhans, S.D., Lienert, J., Schuwirth, N.: The conceptual foundation of environmental decision support. J. Environ. Manage. **154**, 316–332 (2015). https://doi.org/10.1016/j.jenvman.2015.01.053
22. Gregory, R., Failing, L., Harstone, M., Long, G., McDaniels, T., Ohlson, D.: Structured Decision Making: A Practical Guide to Environmental Management Choices. Wiley, Hoboken (2012)