

A Feasibility Study of a Method for Identification and Modelling of Cybersecurity Risks in the Context of Smart Power Grids

Aida Omerovic¹, Hanne Vefsnmo², Gencer Erdogan¹, Oddbjørn Gjerde²,
Eivind Gramme³ and Stig Simonsen³

¹*SINTEF Digital, Norway*

²*SINTEF Energy, Norway*

³*Skagerak Nett, Norway*

Keywords: Cybersecurity, Cyber Risk, Smart Power Grids, Risk Identification, Risk Analysis, Vulnerabilities.

Abstract: Power grids are undergoing a digital transformation and are therefore becoming increasingly complex. As a result of this they are also becoming vulnerable in new ways. With this development come also numerous risks. Cybersecurity is therefore becoming crucial for ensuring resilience of this infrastructure which is critical to safety of humans and societies. Risk analysis of cybersecurity in the context of smart power grids is, however, particularly demanding due to its interdisciplinary nature, including domains such as digital security, the energy domain, power networks, the numerous control systems involved, and the human in the loop. This poses special requirements to cybersecurity risk identification within smart power grids, which challenge the existing state-of-the-art. This paper proposes a customized four-step approach to identification and modelling of cybersecurity risks in the context of smart power grids. The aim is that the risk model can be presented to decision makers in a suitable interface, thereby serving as a useful support for planning, design and operation of smart power grids. The approach applied in this study is based on parts of the "CORAS" method for model-based risk analysis. The paper also reports on results and experiences from applying the approach in a realistic industrial case with a distribution system operator (DSO) responsible for hosting a pilot installation of the self-healing functionality within a power distribution grid. The evaluation indicates that the approach can be applied in a realistic setting to identify cybersecurity risks. The experiences from the case study moreover show that the presented approach is, to a large degree, well suited for its intended purpose, but it also points to areas in need for improvement and further evaluation.

1 INTRODUCTION

Advanced and innovative capabilities are steadily emerging and being deployed on the top of the traditional power grids. Such modern power grids are often called smart grids. New kinds of software and hardware technologies are enablers while increased needs for power grid efficiency are the driving forces for this development, which is characterized as power grid digitalization. With this development come also numerous cybersecurity risks that are more or less specific to complex cyber-physical systems which include many dependencies.

The smart grid vision implies extensive use of "ICT", i.e. information and communication technology, in the power system, enabling increased flexibility and functionality and thereby meeting future demands and strategic goals. Consequently,

power system reliability will increasingly depend on ICT components and systems (Tøndel et al., 2017). While adding functionality, ICT systems also contribute to failures. To analyse the risks of this complex and tightly integrated cyber-physical power system, there is a need to identify the new vulnerabilities that are introduced due to the increasing usage of ICT technologies and their interdependencies with the physical power grid.

The digitalization of the power system will include new concepts based on intelligent sensors in the grid and efficient communication between these sensors and the Supervisory Control and Data Acquisition (SCADA) system or distribution management system (DMS) (Belmans, 2012). New components and technologies, such as self-healing grids, will enable automation of the power grid, which will lead to reduced time for fault- and

interruption handling, but will at the same time introduce new vulnerabilities and threat scenarios, leading to unwanted incidents. One example of how adversaries can exploit the new components and technologies, is the cyber-attack against the Ukrainian Power Grid in December 2015, where the outages affected approximately 225 000 customers that lost power across various areas (Lee et al., 2016). The new architectures of the ICT dominated power system will increase the complexity and therefore calls for approaches to identify the cybersecurity risks of the complex cyber-physical power system.






Risk modelling is a technique for risk identification and assessment, and the state-of-the-art offers several tree-based and graph-based notations. Fault Tree Analysis (IEC, 1990), Event Tree Analysis (IEC, 1995) and Attack Trees (Schneier, 1999) are examples of the former and provide support for reasoning about the sources and consequences of unwanted incidents, as well as their likelihoods. Cause-Consequence Analysis (Nielsen, 1971), CORAS (Lund et al., 2010), and Bayesian networks (Ben-Gal, 2008) are examples of graph-based notations. CORAS is a tool-supported and model-driven approach to risk analysis that is based on the ISO 31000 risk management standard (ISO, 2009). It uses diagrams as a means for communication, evaluation and assessment. Markov models (IEC, 2006), CRAMM (Barber and Davey, 1992), OCTAVE (Alberts et al., 2003), Threat Modelling (Microsoft, 2018) and a number of others, have also been applied to support risk analysis. A framework for studying vulnerabilities and risk in the electricity supply, based on the bow-tie model, has been developed and is published for instance in (Kjølle and Gjerde, 2015, Kjølle and Gjerde, 2012, Hofmann et al., 2012). As stated in (Tøndel et al., 2017) the current methods for risk analysis of power systems seem unable to take into account the full array of intentional and accidental threats. In addition, they found few methods and publications on identification of interdependencies between the ICT and power system (Tøndel et al., 2017). However, as they stand, none of the existing approaches provides the support that meets specific needs for cybersecurity risk analysis of smart power grids. Power grids are namely characterized by very high complexity and a significant degree of interdependencies. As a critical infrastructure which is undergoing a rapid digital transformation, these cyber-physical systems are safety critical and their cybersecurity is becoming crucial. They include assets such as physical power network, communication protocols, control systems, human in the loop, and many emerging types of

hardware and software such as sensors, remote decision support systems, algorithms for automatic response to failures, load balancing, etc. These assets constitute the building blocks for the on-going digitalization efforts, as a means for enabling the power grids of meeting the future needs in terms of capacity, efficiency and reliability. Many of the solutions are new to the domain and there is a lack of formerly established experiences with regard to their strengths and weaknesses. The complexity and the lack of prior empirical knowledge contribute to the inherent uncertainty and the overall risk picture. Moreover, the interdisciplinary nature of such systems poses requirements on comprehensibility of the design of smart power grids and the corresponding risk models. Since the emerging solutions are at their early stages, there is a lack of historical data and operational experiences that could constitute relevant input to the risk models. Even if such data exists, it is of a limited scope and precision in the context of smart power grids. A smart grid setting which includes a complex and critical cyber physical system, human in the loop, uncertainty due to lack of knowledge, many dependencies and interdisciplinary aspects, challenges the state-of-the-art on cybersecurity management within the power distribution sector. This indicates a need for an approach to cybersecurity risk identification which is customized to meet the following **requirements** (the ordering is arbitrary and does not express the relative importance of the requirements):

1. The approach is cost-effective and light-weight, i.e. the benefits of using it are well worth the effort.
2. The cyber risk model can be developed and easily understood by the involved actors who represent varying roles and background.
3. The risk model has sufficient expressive power to capture relevant aspects of the cybersecurity risk picture in the context of smart power grids.
4. The risk model facilitates inclusion of the information that is available, while not requesting unrealistic degree of precision.
5. The risk model can visualize the cybersecurity relevant dependencies and sequence of states/events both for the whole context and for the detailed parts of the scope of analysis.

This paper proposes a customized four-step approach to identification of cybersecurity risks in the context of smart power grids. The aim is that the risk model can be presented to human in a suitable interface, thereby serving as a useful support for decision making during design and operation.

Table 1: Constructs of the cybersecurity risk modelling language.

Symbol	Name	Simple explanation
 Text	Direct asset	Something of a value to the stakeholders and needs to be protected. The risks are to be identified with respect to the direct assets.
 Text	Vulnerability	A state or a property that may be exploited.
 Text	Threat scenario	A relevant event or a property that does not directly harm an asset.
 Text	Unwanted incident	An event that directly harms an asset.
 Text	Indirect asset	Something of a value to the stakeholders and can be affected by a direct asset.

The approach applied in this study is, to a high degree, based on parts of the previously mentioned CORAS method for model-based risk analysis. Compared to CORAS, the process and the modelling approach we have applied are simplified and partially adapted in order to meet the five above specified requirements.

The paper also reports on results and experiences from applying the approach in a realistic industrial case study together with a power distribution system operator (DSO) responsible for hosting a pilot installation of the self-healing functionality within a medium voltage (MV) grid. The results indicate feasibility and usefulness of the approach. Important lessons have been learned on our approach to risk identification as well as about the cybersecurity of the power grid domain. This work has been carried out within the CINELDI research centre (CINELDI, 2018) that performs research on the future intelligent energy distribution grids. The centre gathers a significant number of the major public and private actors from the energy sector in Norway.

The rest of this paper is organized as follows: In Section 2 we briefly present the research strategy applied. Our method for cybersecurity risk identification is presented in Section 3. The setup and the results from the trial of the method are outlined in Section 4. In Section 5 we discuss the results, before concluding in Section 6.

2 RESEARCH STRATEGY

Figure 1 illustrates the steps of our research strategy, which is in line with the design science approach (Wieringa, 2014). Although Figure 1 illustrates sequential steps, the method was carried out

iteratively where some of the steps were revisited during the process.

In Step 1, we identified the five requirements for a model-based risk identification approach customized for smart power grids based on the background and needs as explained in Section 1. In Step 2, we developed the risk identification approach based on state-of-the-art approaches with respect to the requirements identified in Step 1. The method consists of four main phases, namely: context establishment, risk identification, risk model validation, and follow-up. The method is explained in detail in Section 3.

Finally, in Step 3 we evaluated the approach in an industrial setting together with a DSO that is currently hosting a pilot on self-healing functionality in their power grid. The evaluation was carried out as follows. First, we established the context and gained a deep understanding of the self-healing functionality, based on reports on state of the practice, dialogue with domain experts who participated in the analysis group, as well as the documentation provided by the DSO. Then, we developed a preliminary version of the risk model which focused on cybersecurity aspects of reliable energy supply. Thereafter, we presented the preliminary version of the risk model to the DSO, i.e. the power distribution company that was the use case provider.



Figure 1: Research strategy.

The company provided their feedback to our risk model. Next, the company presented results of a risk

assessment that they had previously carried out. Their assessment had focused on preservation of human safety in the context of self-healing. Both the feedback and the risk assessment of the company were used as a basis to revise our risk model. This feedback-correction interaction was carried out in several iterations. This was followed by a final revision that involved updating the model according to the predefined scope. As the final step of the evaluation, we exposed the model to a list of future scenarios that had independently been developed in two workshops by industry experts in the CINELDI research centre. The complete evaluation process is presented in Section 4.

3 METHOD FOR CYBERSECURITY RISK IDENTIFICATION – THE MODELLING APPROACH AND THE PROCESS

Our method for cybersecurity risk identification consists of four main phases, inspired by the CORAS method and modelling language, but simplified and customized in order to address our specific requirements listed in Section 1. The method is to be carried out by an analysis group, consisting of analysts and domain experts, representing competence in risk analysis, cybersecurity, and smart power grids. One individual may cover one or more roles, and several individuals may represent a similar role. Most importantly, the composition of the analysis group needs to include the relevant competence and ensure a sufficient degree of continuity (with respect to attendance of some of the participants) within the group. The four phases of the process include:

1. Context establishment
2. Risk identification
3. Risk model validation
4. Follow-up

The objective of Phase 1 is to characterize the scope and the target of the analysis. Stakeholders that the analysis is being performed on behalf of, time perspective, relevant terminology, assumptions, roles and participants of the analysis group, as well as the information sources are, moreover, specified. Assets, that is the values that will drive the focus of the analysis, are also defined. This phase produces descriptions, insights and a common understanding of the target of analysis. The target of the analysis is

specified and modelled with respect to capabilities, structure, dataflow, workflow, etc. The existing target specifications (i.e. those which are available prior to the analysis) may be reused or referred to.

The objective of Phase 2 is to identify the relevant risk model elements and develop a risk model. The risk model elements may be of the following types: vulnerabilities, threat scenarios and unwanted incidents. The unwanted incidents are the elements that may harm the assets defined during Phase 1. The very first step of this phase is to introduce the types of the model elements to the analysis group. A brief introduction to the modelling constructs and their simple explanations is illustrated in Table 1. The explanations in the last column are simplified wordings inspired by corresponding definitions from CORAS. Note that the definition of vulnerability from the energy sector is slightly different, namely "*Vulnerability is an expression for the problems a system faces to maintain its function if a threat leads to an unwanted event and the problems the system faces to resume its activities after the event occurred. Vulnerability is an internal characteristic of the system*" (Kjølle and Gjerde, 2015). Thereafter, the identification of risks through a risk modelling activity using the constructs in Table 1, is initiated. The constructs are, as a part of this process, annotated with descriptive text. The relationships between the model constructs are expressed with arcs connecting the relevant elements, thus resulting in a risk model shaped as an acyclic directed graph. The analysts shall facilitate the model development by iteratively posing questions on risks that may harm the assets and the possible vulnerabilities and threat scenarios that cause those risks. The analysts shall also contribute with cybersecurity domain knowledge during the risk modelling. The analysts shall, moreover, ensure that the syntax of the risk model is consistent. The domain experts shall, during the risk modelling, contribute with the domain knowledge on power grids. Discussion is facilitated in order to align the different domains and reveal the relevant risks- At the same time, the context description from Phase 1 is actively used. Moreover, if needed, refined descriptions of selected model elements are provided. For some parts of the risk model, it may be appropriate to express uncertainties and assumptions, in form of supplementary information or within the model.

Phase 3 aims at validating the risk model developed in the preceding phase. That is, the model should be exposed to quality assurance based on various and complementing kinds of empirical input, in order to ensure an acceptable level of uncertainty.

The uncertainty may origin from insufficient information or knowledge, or from variability in context, usage, etc. This is followed by adjustments of the model with respect to the structure and the individual elements. Eventually, the model is approved if the evaluation shows that the revised version is sufficiently complete, correct and certain.

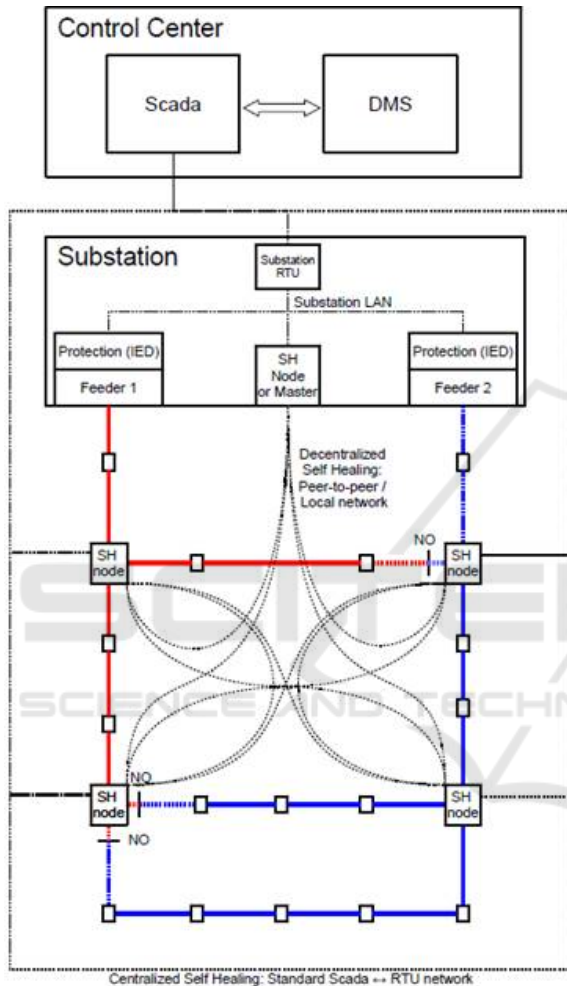


Figure 2: The centralized self-healing design network. SH-nodes are the self-healing nodes consisting of fault indicators and remotely controlled switches.

The objective of Phase 4 is two-fold, namely communication and maintenance of the results. The specific tasks of this phase include summary of most critical findings, evaluations of validity and reliability of the risk model, recommendations of risk treatments, summary of uncertainties in the findings, communications of the results to the relevant stakeholders. Maintenance of the risk model involves monitoring of assumptions and the context changes that require updates of the risk model.

4 TRIAL OF THE METHOD ON AN INDUSTRY PILOT WITH SELF-HEALING CAPABILITY

This section outlines the process undergone during the industrial case, as well as the main properties and examples of the modelling artefacts produced. We also summarize the lessons learned. Note that the process undergone is to a large degree but not fully an instantiation of the approach presented in Section 3. The reason is that the process has been simplified in order to meet the most prevailing needs and to apply as much as possible of the approach, within the limited resources assigned.

4.1 Setting of the Case Study

New sensors and technologies give opportunities for smarter fault- and interruption handling in the distribution (medium voltage) grid through self-healing grid functionality. At the same time these technologies introduce new vulnerabilities to the system. To study these vulnerabilities a case study has been performed on detailed design of a realistic centralized self-healing grid, owned by a Norwegian DSO.

4.1.1 Centralized Self-healing Grid

With self-healing functionality the processes as fault location, isolation and restoration (FLIR) (Siirto, 2016) are, in our context, assumed to be fully automated. Centralized self-healing systems are traditionally implemented in the control centre, but in this case the logic and algorithm are implemented in the "Master node" in the substation, as shown in Figure 2. This Master node collects the information from all self-healing nodes ("SH-nodes" in Figure 2) and makes the best-effort decision based on this information. The SH-nodes consist of fault indicators and remotely controlled switches. The white rectangles are nodes with load points. The red and blue lines indicate the two feeders which are radially operated power lines from the actual substation. The "NO" in Figure 2 shows the shifts from feeder 1 and feeder 2 and by that which loads are supplied from which feeder. The Master node communicates with the substation remote terminal unit (RTU), which again communicates with the control centre. The centralized self-healing system uses the same communication network as SCADA/DMS to communicate with remote devices such as substation communication. As the system is normally inside the same security zone as the control centre and the

remote devices are normally on the outside, communication must pass the barriers in and out of the security zone (Tutvedt et al., 2017).

4.1.2 Scope of the Analysis

The focus of the analysis has been to protect the following two assets: reliability of the electric power system and human safety. The reliability of an electric power system is defined as: "*probability that an electric power system can perform a required function under given conditions for a given time interval.*" (IEV 617-01-01).

When a fault occurs in a radially operated medium voltage grid, all customers experience an interruption. The goal of self-healing is not to avoid faults, but to reduce the consequence in terms of interruption duration, which is obtained when self-healing works perfectly after a fault. When self-healing functionality fails, the interruption duration increases and may also be longer than without any self-healing solution. In this case study, risks leading to longer interruption durations are identified and modelled, together with new vulnerabilities which may introduce new fault types into the power system.

In addition, safety risks caused by security risk are included in the analysis. This may cover safety aspects related to personnel working close to the power line or public safety, in order to ensure nobody gets hurt during power restoration.

4.1.3 Assumptions and Delimitation of the Industrial Case

The following assumptions were made in the case study as part of context establishment:

- The self-healing starts after a fault has occurred.
- Self-healing session terminates when the fault is isolated, and the power is resupplied to all possible end-users. The repair of the fault is not a part of the scope of the analysis.
- Self-healing in the distribution grid is still at an early stage and under implementation. The centralized self-healing analysed is therefore based on the version/design available at the time of case execution.
- Configuration of self-healing and fault handling of the self-healing system (i.e. if the self-healing functionality fails) is included in the scope.
- It is assumed that self-healing is performed autonomously (That is, automatically without an operator involved) when the self-healing functionality works as intended.

- The operator is involved in the configuration and the fault-handling part, i.e. if the self-healing functionality fails.

4.1.4 Background of the Participants Involved in the Analysis

The analysis was performed with a core team of four people. The main analysis was performed by the two analysts with 8-16 years of experience in software engineering and cybersecurity risk management. In addition, two domain experts with 6-18 years of relevant experience within power system reliability and security of electricity supply, have participated. From the grid company, one manager and two domain experts within grid operation and development, have been involved. The manager has 3 years of experience from the grid company and additional 9 years of experience from a technology provider in the power system. The operation and grid development experts have more than 10 years of relevant experience from power grid planning and operation.

4.2 Process Outline

The case study was conducted during the second half of the year 2018. The analysis was performed in the form of one physical and eight videoconference meetings in a fully realistic setting in terms of the scope, the objectives, the process, the risk models and the participants. The cybersecurity risks of the self-healing pilot were identified and modelled, with respect to a specified scope and assets. The case study was conducted first as a part of risk modelling with context specification as a baseline, and then as a part of method evaluation with an independent risk analysis previously conducted by the power grid operator as a baseline.

Table 2 summarizes the process undergone. For each workshop, we list the meeting number, the date, the participants, the meeting type, the meeting length, and activities.

4.3 Results from the Case Study

Figure 3 illustrates a high-level view of the risk model developed in the study. In the following we first explain the textual notation used in the model as shown in Figure 3 and then we explain the content of the risk model by referring to its elements.

The risk model in Figure 3 has 27 threat scenarios (*TS_01 – TS_27*), seven unwanted incidents (*UI_01 – UI_07*), two assets (*A1* and *A2*) and one indirect asset (*A3*). The digit in each vulnerability indicates the

Table 2: The process undergone during the case.

Meeting	Activities
No. 1; Date: 01.06.2018; Participants: 1 analyst, 1 domain expert; Meeting type: physical; Duration: 1,5h;	Identification of context, goals, scope and focus for the case study. Clarification of main concepts in self-healing, and the asset to consider in risk identification. Walk-through of an example and parts of references that previously had been made available. Preparations: The analyst had prepared a set of questions based on studies of the material that previously had been made available.
No. 2; Date: 29.06.2018; Participants: 2 analysts, 2 domain experts; Meeting type: video; Duration: 3h;	The analysts presented their initial understanding of the context, the initial version of the risk model, as well as a template that structures presentation of the context, the process and the main results. The risk model was discussed and updated.
No. 3; Date: 15.08.2018; Participants: 2 analysts, 3 domain experts; Meeting type: video; Duration: 2h;	The risk model was discussed and updated. Planning of next steps, including a case-based evaluation.
No. 4; Date: 29.08.2018; Participants: 1 analyst, 2 domain experts; Meeting type: video; Duration: 1h;	The risk model was discussed and updated. Detailed planning of next steps, including a Case-based evaluation.
No. 5; Date: 14.09.2018; Participants: 2 analysts, 4 domain experts (incl. 2 from industry); Meeting type: video; Duration: 2h;	The current version of the risk model was presented to the industry partner. Feedback was received on the general impression, as compared to the results of a risk identification of self-healing that the grid company itself had previously performed.
No. 6; Date: 09.10.2018; Participants: 2 analysts, 4 domain experts (incl. 2 from industry), 1 manager from industry; Meeting type: video; Duration: 1h;	The grid company presented the results of their own risk analysis of self-healing, as well as the underlying context. Parts of the documentation from this analysis were also made available to the analysis team from SINTEF, so that our risk model could be complemented.
No. 7; Date: 19.10.2018; Participants: 2 analysts, 4 domain experts (incl. 2 from industry), 1 manager from industry; Meeting type: video; Duration: 1h;	The grid company presented the results of their own risk analysis of self-healing, as well as the underlying context. Parts of the documentation from this analysis were also made available to the analysis team from SINTEF, so that our risk model could be complemented.
No. 8; Date: 08.11.2018; Participants: 1 analyst, 1 domain expert; Meeting type: video; Duration: 1,5h;	The risk model was discussed and updated.
No. 9; Date: 12.11.2018; Participants: 1 analyst, 1 domain expert; Meeting type: video; Duration: 1h;	The risk model was discussed and updated into a final version.

number of vulnerabilities that may lead to one or more threat scenarios and/or unwanted incidents. For example, in the top-left corner of Figure 3, we see one vulnerability that has the digit 9 conveying that there are nine different vulnerabilities that may only lead to threat scenario *TS_02*. Moreover, we see that there is one vulnerability that may lead to both threat scenarios *TS_02* and *TS_05*. Thus, the risk model in Figure 3 shows that there our final risk model contained in total 62 vulnerabilities.

Some of the threat scenarios and unwanted incidents are closely related and therefore aggregated and illustrated by a dashed rectangle. For example, the aggregated threat scenario *A_TS_01* consists of threat scenarios *TS_02*, *TS_03* and *TS_04*. The model also shows an aggregation of unwanted incidents that

are related (*A_UI_01*). In the following, we explain the model by referring to its elements in Figure 3.

With respect to assets to protect, the model captures the following:

- *A_1*: Human safety.
- *A_2*: Reliability of electricity supply.
- *A_3*: Security of other critical infrastructure.

With respect to threat scenarios, the model captures the following:

- *A_TS_01*: Malicious users access SCADA via underlying infrastructure or due to misconfiguration of the power grid and carries out cyber-attacks on services.

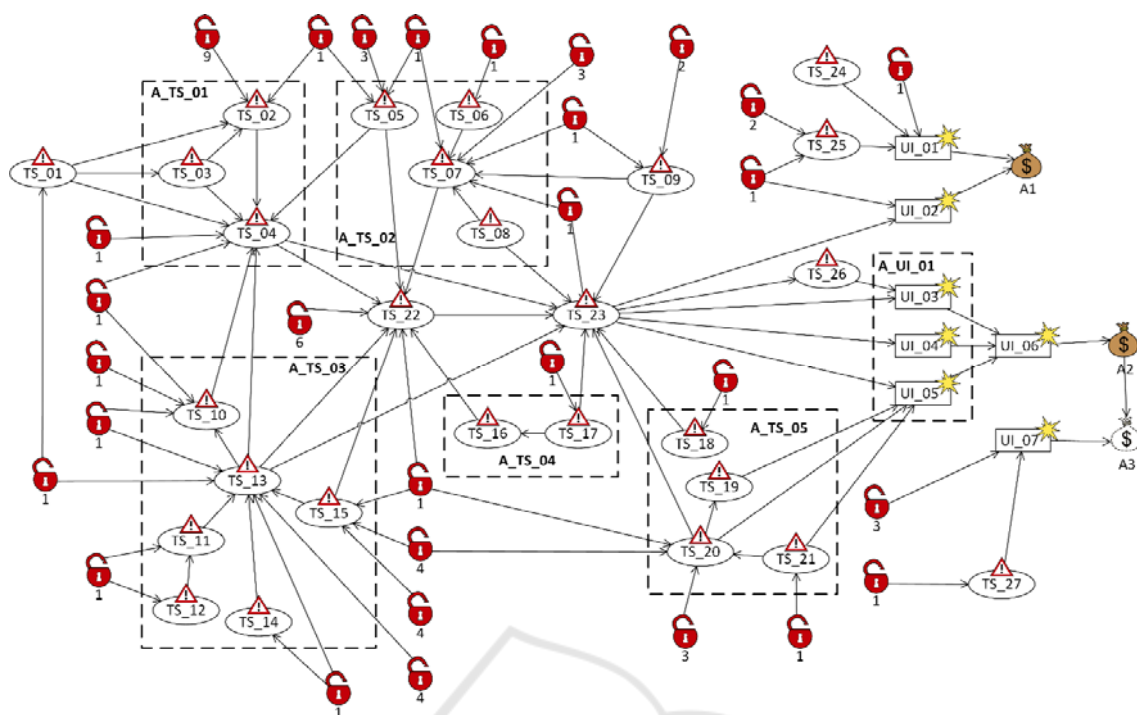


Figure 3: High level view of the risk model. *TS* means Threat Scenario. *UI* means Unwanted Incident. *A* means Asset. *A_TS* means Aggregated Threat Scenario.

- *A_TS_02*: The human operators and/or the autonomous workflows misconfigures the grid based on erroneous information.
 - *A_TS_03*: Self-healing nodes produce erroneous information due to failure of sensors and communication channels or maliciously altered sensor data.
 - *A_TS_04*: Outdated information in the SCADA system prohibits the operators in getting insight into whether self-healing is taking place due to disruption of communication between the centralized communication unit and the SCADA system.
 - *A_TS_05*: The protection/breaker is exposed to missing or unwanted tripping due to technical issues, lack of power supply from batteries, insufficient maintenance, or wear and tear.
 - *TS_01*: Unauthorized user accesses the communications network from associated infrastructure and uses that as a backdoor to access the smart power grid.
 - *TS_22*: The decision support system produces inadequate or misleading information.
 - *TS_23*: The switch links are erroneous or delayed.
 - *TS_24*: Automatic remote control causes a complex situation.
 - *TS_25*: Unintentional voltage setting.
 - *TS_26*: The self-healing connection sequence does not lead to isolation of the erroneous location.
 - *TS_27*: Unauthorized user exploits granted access to parts of the smart power grid as a backdoor.
- With respect to unwanted incidents, the model captures the following:
- *A_UI_01*: Delayed or prevented sectioning, or sectioning of incorrect area.
 - *UI_01*: Personnel are not safe while operating in self-healing network.
 - *UI_02*: The switching system is connected automatically or remotely while there is ongoing work on the grid.
 - *UI_06*: Prolonged duration of interruption.
 - *UI_07*: Unauthorized user gains access to smart grid and further to other critical infrastructure.
- Due to confidentiality aspects, the power grid company involved in the case requested not to reveal all details of the risk model. However, we can illustrate a fragment of the model considering the vulnerabilities. Figure 4 illustrates the threat scenario *TS_02* and the nine vulnerabilities that only lead to *TS_02*.

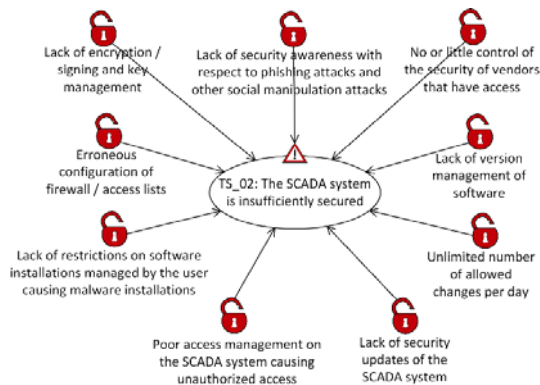


Figure 4: The threat scenario TS_02 and the nine vulnerabilities that only lead to TS_02.

4.4 Experiences and Lessons Learned

In the following, we summarize the experiences and the lessons learned from applying our approach on the aforementioned self-healing case.

A thorough understanding of the context is crucial. In fact, a deep understanding of the self-healing capability was gained by all analysis participants including the analysts, which also enabled an active participation in risk modelling and development of the preliminary model by the analysts. Such a deep insight into the domain is not assumed by the method, but was in our case experienced as beneficial, in spite of the additional resources spent on that. It is generally crucial to dedicate sufficient resources to context establishment.

Regardless of how well the security risk analyst understands the context, it is crucial that the analyst does not develop the risk models alone. Risk identification triggers many useful discussions among the analysis participants, and helps reveal inconsistencies and misunderstandings. In addition, risk models are intended to be used by the industry and the domain experts, who need to be able to maintain the models. The optimal approach, in our point of view, is that the analyst presents an initial version of the risk model, which is then discussed, corrected and further developed in the group. Errors or missing parts in the initial models are often an advantage, as they trigger the discussions in the group.

The analyst has to be aware of the inconsistencies of the terminology used in documents and the verbal communication among the domain experts, as well as between the overall stakeholders. Any such inconsistencies should be clarified.

While the power grid company had considered the

incidents believed to be of less frequent nature, our original model covered the incidents and vulnerabilities of in principle more "known" nature.

While the grid company had applied a bottom-up approach to risk identification in the sense that they started at a lower level of abstraction driven by the various threat scenarios associated with the new capabilities in self-healing, our approach was of a rather top-down nature and driven by the pre-specified asset.

Being a driving force of the risk identification, assets showed to be very beneficial in order to keep focus on the relevant aspects of the risk model and ensure the right abstraction level. While the grid company had focused on safety, our asset was reliability of electricity supply. This made it possible to complement the originally separate results into a merged and comprehensive model.

An important property of the model was that it was possible to express all information into one merged model, so that the complexity and all relevant relationships were explicit and fully reflected in a single comprehensive overview.

Our original model had explicitly focused on the cybersecurity risks that could impact the reliability of electric supply, while the risk model of the grid company had addressed the risks of more general type that could impact the safety. The final merged model included all relevant aspects and showed how the cybersecurity risks affect the assets concerning not only the resilience of the power grid and the systems associated, but also the human safety. This has enabled us to explicitly visualize the dependencies between the several types of risks, on all assets. As such, the risk model has bridged the gap between cybersecurity and safety within our context.

Several of the analysis participants did not have prior experiences in either CORAS or graphical risk modelling in general. Still, the risk model was gradually updated into new versions through iterative and thorough discussions among all the participants. Our observation is that the model offered an adequate level of abstraction that made it easy to understand and contribute to.

Through the final risk model, we were able to express all cybersecurity relevant risks and dependencies among various risk elements that were suggested by the analysis group, either based on own knowledge, or based on the references provided through context establishment.

The grid company expressed that the graphical risk modelling was appropriate and enabled the model to be significantly more structured and comprehensible.

Prior to and fully independently of our case study, a set of so-called future scenarios, i.e. various contexts that illustrate characteristics associated with power grids of the future, had been brainstormed and described through a workshop. The workshop participants were system operators, technology providers and researchers, all from the energy domain. The future scenarios were, after our case study, used for the purpose of the quality assurance of the final risk model. At this stage, no further modifications of the model were found necessary.

5 DISCUSSION

Based on our experience, we discuss and evaluate in this section the fulfilment of the requirements one through five defined in Section 1. The second part of this section discusses the main threats to validity and reliability of the results.

5.1 To What Degree Are the Requirements Fulfilled?

Requirement 1: Our estimate based on the case study indicates that the method for cybersecurity risk identification described in this paper amounts to ca. 150 man-hours in total spent by the analysts. This includes ca. 23 hours spent on meetings 1 to 9 as described in Section 4.2. Ca. 127 man-hours were spent on work between the meetings. That is, hours spent by the analyst team before and after each meeting including the process of establishing the context, identifying risks, preparing the meetings, and taking notes and correcting the model during and after the meetings. The authors of the CORAS risk analysis method state that the expected effort required to carry out a CORAS analysis is typically from 150 to 300 hours (Lund et al., 2010), while in our case we spent in total ca. 150 hours. This indicated that the amount spent on carrying out the approach outlined in this paper is reasonable. It should also be taken into account that our analysts gained such a deep understanding of the context, that they were enabled to actively identify and model risks. This is not a general assumption within the above mentioned budgets. Our approach is light-weight in the sense that it does not contain steps such as risk estimation, evaluation and treatment, which are typically found in full scale risk analysis methods. This also removes the overhead in the context establishment phase where, in our case, it is not necessary to define likelihood and consequence scales often used to estimate and evaluate risks. Of course, these are steps

that are necessary when there exists data to support risk estimation and when the goal is to estimate and evaluate risks. However, as pointed out in Section 1, self-healing within smart power grids is a state-of-the-art approach of limited maturity and empirical evidence, which would be needed for risk estimation. To fully justify Requirement 1, we would have to quantify the benefit, as well as the cost. This is very hard, and we have not attempted to do so. However, the feedback received and the experiences indicate that the benefit justifies the effort, meaning that our customized method is reasonably cost effective as well as light-weight.

Requirement 2: As outlined in Section 4.2, the participants vary with respect to background. The domain experts were experts in smart power grids and had barely limited background in cybersecurity. Even so, after a brief introduction of the cyber risk concepts explained in Section 3, they quickly grasped the concepts and were able to actively contribute to the model development. Moreover, the comments, suggestions and discussions throughout the process demonstrated that all participants were able to understand the details of the evolving model. This is further substantiated by earlier studies which have empirically shown that the graphical notation used in our method is intuitively simple for stakeholders with very different backgrounds (Volden-Freberg and Erdogan, 2019, Solhaug and Stølen, 2013).

Requirement 3: Prior to the study reported in this paper, it was not clear whether the modelling language used in our method had sufficient expressiveness in terms of capturing relevant risks of smart power grids. As part of the study, we were able to capture all vulnerabilities, threat scenarios, unwanted incidents and assets identified by the domain experts. Even if the various threat scenarios, vulnerabilities, unwanted incidents and assets were sometimes different in nature (technical versus business risks), no relevant risks were left out. This was pointed out by the case providers (the DSO) to be an advantage because it helped understanding the relationship between high-level business risks and low-level technical risks. It is therefore reasonable to argue that our risk modelling approach has sufficient expressiveness to capture relevant aspects of the cybersecurity risk picture in the context of smart power grids.

Requirement 4: As discussed in the previous sections, self-healing smart power grids are still immature and advancing. Thus, there is little experience in their usage including operational data. There is also very little experience in terms of cyber risk incidents in this context. Due to these factors, we

have deliberately designed our method to facilitate a rather simple modelling language in which only the necessary risk-related concepts are used: threat scenarios, vulnerabilities, unwanted incidents and assets to protect. As illustrated in Figure 3 and Figure 4, and explained in Section 3, these risk model elements are left to the analyst to describe at a qualitative level. No unavailable information of unrealistic precision is demanded, unless possible and desired to include. The analyst is therefore free to choose the level of abstraction in which the threat scenarios, vulnerabilities, and their relationships, are described. This flexibility eases the risk identification process and does not "lock" the analysts to think at one level of abstraction (for example, identifying risks only at technical level or only at business level, or having to deal with quantitative risk estimates).

Requirement 5: As illustrated in Figure 3 and figure 4, the risk model produced by our method is capable of illustrating the risk picture both for the whole context and the detailed parts of the scope of analysis.

The risk model in Figure 3 illustrates a high-level view of the risk picture in terms of threat scenarios (including aggregated threat scenarios and aggregated unwanted incidents), unwanted incidents caused by threat scenarios, assets harmed, and the number of vulnerabilities associated with the threat scenarios and unwanted incidents. We could have chosen to provide unique IDs for each vulnerability and illustrate all in the high-level risk model in Figure 3. In that case, we would have had to add 62 vulnerability icons instead of the 30 shown. However, to support scalability, we chose to aggregate the vulnerabilities in the high-level risk model by only illustrating the number of vulnerabilities associated with the various elements of the risk model, as well as illustrating aggregated threat scenarios and aggregated unwanted incidents as explained in Section 4.3. Figure 4, on the other hand, illustrates a fragment of the detailed risk model in which we can see all risk model elements including their description.

The advantage of the high-level risk model in Figure 3 is that the analysts as well as the domain experts get a top-down view which helps obtaining an overall risk picture by a quick glance. This helps identifying the most vulnerable parts of the target system, as well as understanding the chain of events in how certain threat scenarios (or group of threat scenarios) lead to other threat scenarios and/or unwanted incidents, which ultimately harms the assets. The advantage of the detailed risk model (partly illustrated in Figure 4), on the other hand, is to get a more refined view of the risk model and obtain detailed description of the various aggregated

vulnerabilities, threat scenarios, and unwanted incidents.

5.2 Threats to Validity and Reliability

Apart from certain parts of the follow-up phase, we have, through the trial, instantiated the entire approach. Application of the approach on a case with limited empirical evidence, has clear limitations in terms of representativeness of the target of the analysis. Still, the results of the trial indicate feasibility of applying the approach. The fact that new knowledge was gained about the target of analysis and its security risks, suggests usefulness of the approach.

Correctness and relevance of the results are partially evaluated through the close interaction with the industry pilot, and through the future scenarios. A retrospective evaluation would have been appropriate, but the industry pilot has not been running for long enough in order to have sufficient retrospective picture of security risks. Instead, we have relied on the analysis group as well as the industry partner, together representing relevant and broad domain knowledge.

It is, in terms of evaluation of performance of the approach, a weakness that the analysts who tried out the approach also participated in design of the customized version of the approach. As such, it is also a threat to reliability of the evaluation results, as we cannot know to what degree another analysis group would have obtained the same results.

There is a need for a baseline for comparing this approach with the alternative risk identification methods, in order to assess characteristics such as usability, scalability, usefulness and cost-effectiveness of our approach compared to the alternative ones. We have already partially compared the evaluation results with the performance of CORAS. Further empirical evaluation in other realistic settings is, however, still needed due to threats to validity and reliability.

Overall, we have drawn important findings and learned lessons from developing this smart-grid-customized approach to cybersecurity risk identification, and from trying it out in the industrial case. Although the mentioned threats to validity and reliability are present in the study, we argue that the results indicate feasibility and usefulness of the approach. Important findings are also obtained on strengths and weaknesses of the approach, and they will guide the directions for our future work.

6 CONCLUSIONS AND FUTURE WORK

Smart power grids are evolving into increasingly complex cyber-physical systems that introduce numerous kinds of interdependencies and vulnerabilities. This poses new requirements to cybersecurity management. As a result, state-of-the-art to risk identification is being challenged. In this paper we have proposed a customized four-step approach for identification and modelling of cybersecurity risks in the context of smart power grids. Our approach is, to a high degree, based on parts of the CORAS method for model-based risk analysis. Compared to CORAS, the process and the modelling approach we have applied are simplified and partially adapted in order to meet the identified requirements. The qualitative nature of the model gave the needed simplicity. The approach has been tried out on a realistic industrial case. The context of the case was a centralized self-healing pilot (i.e. a limited pilot installation of self-healing functionality within a medium voltage power distribution grid). We argue that our approach to some extent fulfils the pre-identified requirements. However, there are at the same time, clear limitations in terms of reliability of the current results, due to e.g. bare evaluation of the approach on one case with limited empirical evidence. Although the mentioned threats to validity and reliability are present in the study, we argue that the results indicate feasibility and usefulness of the approach.

The next step will be to test out the developed approach on another case study in the smart grid domain. This will give us the opportunity to further analyse the scalability, performance, relevance and representativeness of our approach. Another next step is to develop recommendations and guidelines for cybersecurity risk identification in the smart grid domain based on our approach. This requires further case studies and adjustments according to our mentioned strengths and weaknesses.

ACKNOWLEDGEMENTS

This paper has been funded by CINELDI - Centre for intelligent electricity distribution, an 8-year Research Centre under the FME-scheme (Centre for Environment-friendly Energy Research, 257626/E20). The authors gratefully acknowledge the financial support from the Research Council of Norway and the CINELDI partners.

REFERENCES

- Alberts, C., Dorofee, A., Stevens, J. and Woody, C. (2003) *Introduction to the OCTAVE Approach*: Carnegie Mellon University.
- Barber, B. and Davey, J. (1992) 'The use of the CCTA risk analysis and management methodology CRAMM in health information systems', *Proceedings of the 7th International Congress on Medical Informatics*, pp. 1589-1593.
- Belmans, R. (2012) *Strategic Research Agenda for Europe's Electricity Networks of the Future - SmartGrids SRA 2035*: European Technology Platform SmartGrids.
- Ben-Gal, I. (2008) 'Bayesian networks', *Encyclopedia of statistics in quality and reliability*, 1, pp. 1-6.
- CINELDI (2018) *CINELDI*. Available at: <https://www.sintef.no/cineldi> (Accessed: December 2018).
- Hofmann, M., Kjølle, G. and Gjerde, O. (2012) 'Development of indicators to monitor vulnerabilities in power systems.', *Proceedings of the 11th International Probabilistic Safety Assessment and Management Conference and the Annual European Safety and Reliability Conference 2012*: Curran Associates, Inc., pp. 5869-5878.
- IEC (1990): *IEC 61025:1990 Fault tree analysis (FTA)*: International Electrotechnical Commission.
- IEC (1995): *IEC 60300-3-9:1995 Dependability management - Part 3: Application guide - Section 9: Risk analysis of technological systems*: International Electrotechnical Commission.
- IEC (2006): *IEC 61165:2006 - Application of Markov techniques*: International Electrotechnical Commission.
- ISO (2009): *ISO 31000: Risk Management - Principles and Guidelines*: Geneva: International Organization for Standardization.
- Kjølle, G. and Gjerde, O. (2012) 'Risk Analysis of Electricity Supply.', *Risk and Interdependencies in Critical Infrastructures: A Guideline for Analysis*: Springer, pp. 95-108.
- Kjølle, G. and Gjerde, O. (2015) 'Vulnerability analysis related to extraordinary events in power systems.', *Proceedings of the 2015 IEEE Eindhoven PowerTech*: IEEE, pp. 1-6.
- Lee, R. M., Assante, M. J. and Conway, T. (2016) *Analysis of the Cyber Attack on the Ukrainian Power Grid: Defense Use Case*, Washington: Electricity - Information Sharing and Analysis Center.
- Lund, M. S., Solhaug, B. and Stølen, K. (2010) *Model-driven risk analysis: the CORAS approach*. Springer.
- Microsoft (2018) *Security Development Lifecycle*. Available at: <https://www.microsoft.com/en-us/SDL> (Accessed: November 2018).
- Nielsen, D. S. (1971) *The cause/consequence diagram method as a basis for quantitative accident analysis*, Roskilde, Denmark: Risø National Laboratory Risø-M, No. 1374).
- Schneier, B. (1999) 'Attack trees: Modeling security threats', *Dr. Dobb's Journal*, 24(12), pp. 21-29.

- Siirto, O. (2016) *Distribution automation and self-healing urban medium voltage networks*.
- Solhaug, B. and Stølen, K. (2013) 'The CORAS Language - Why it is designed the way it is', *Proceedings of the 11th International Conference on Structural Safety and Reliability*: Taylor and Francis, pp. 3155-3162.
- Tutvedt, K. A., Seguin, R., Kjølle, G., Simonsen, S., Hermansen, T. S. and Myhr, I. (2017) 'Smart fault handling in medium-voltage distribution grids', *CIREC-Open Access Proceedings Journal*, 2017(1), pp. 1471-1474.
- Tøndel, I. A., Foros, J., Kilskar, S. S., Hokstad, P. and Jaatun, M. G. (2017) 'Interdependencies and reliability in the combined ICT and power system: An overview of current research', *Applied Computing and Informatics*, 14(1), pp. 17-27.
- Volden-Freberg, V. and Erdogan, G. (2019) 'An Empirical Study on the Comprehensibility of Graphical Security Risk Models Based on Sequence Diagrams', *Proceedings of the 13th International Conference on Risks and Security of Internet and Systems*: Springer, pp. (To Appear).
- Wieringa, R. J. (2014) *Design science methodology for information systems and software engineering*. Springer.

