DETERMINING FAILURE PROBABILITIES WITH QUALITATIVE CHARACTERISTICS FOR RISK ESTIMATION IN THE SGAM

Marie VAN AMELSVOORT
OFFIS Institute for Information Technology – Germany
Marie.vanamelsvoort@offis.de

ABSTRACT

In this work, we address the subject of Security-by-Design in Smart Grids. Our approach bases on designing technologies with the Smart Grid Architecture Model and existing risk estimation formulas, where we will look at characteristics to help estimating failure probabilities. Afterwards we will discuss which way of combining these characteristics and failure probabilities fits best for a risk estimation and how it can be improved.

INTRODUCTION

Renewable energies are an advancing technology and help to reduce CO_2-emissions. Smart grid technologies were developed to modernize and improve traditional electrical grids to become smart grids. So here we have a strong evolving field and furthermore, the energy supply is a critical infrastructure in everyday-life. Taking together the points of evolvement and criticality, it is desirable that new Smart Grid technologies are secure – and therefore security-by-design is needed.

In this paper we will analyze an approach for risk estimation in design-time. Therefore, we take the well-established Smart Grid Architecture Model (SGAM) to design and model Smart Grid technologies and continue analyzing how to integrate the viewpoint *security* in this model, where our approach attaches the analysis of security levels for the SGAM cells.

Methodical Design Approach

In regard of security-by-design, we will first introduce the corresponding design methodology. The idea is a complete toolchain from the description of a use case up to a first rating for the described technology. This toolchain is depicted in Figure 1and starts, after the idea and conception for a new technology, with a description of the corresponding Use Case. This description can be made with the Use Case template provided by the Public Available Standard IEC PAS 62559 [1]. This Use Case description can be imported within a Use Case Management Repository (UCMR) and from the informations within the Use Case description, the architectural design of the technology can be described within the Smart Grid Architecture Model, that will be explained in the next section. Having the architectural description, the value of Key Performance Indicators (KPI) can be estimated so that a first rating of the technology can be made and to compare it with other technologies. The overall-rating of technologies descripted with the Smart Grid Architecture Model is further described in [2].

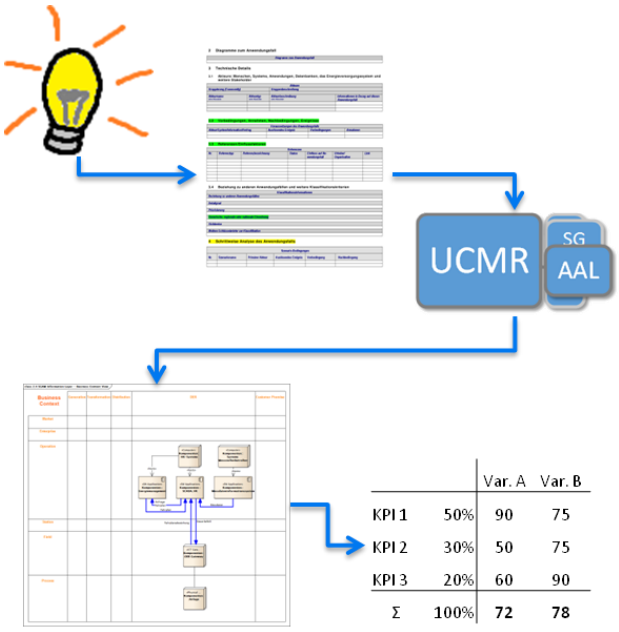

Figure 1: Toolchain from Use Case description to rating key performance indicators

In this work, we will focus on the aspect of security within this rating integrated in the toolchain. The goal is to find KPIs and a corresponding metric to derive a risk estimation for technologies.

Smart Grid Architecture Model

The Smart Grid Architecture Model (SGAM) was developed in the context of the European Commission's standardization mandate M/490 to CEN, CENELEC and ETSI and is introduced in [3]. It provides a holistic viewpoint of an overall Smart Grid infrastructure and can be used to design and model Smart Grid technologies.

As depicted in Figure 2, the SGAM-plane consists of five domains referring to the supply chain in the energy sector and six zones according to the hierarchical energy management zones. This SGAM-plane builds the basis structure for five interoperability layers, which were adopted from the GridWise Architecture Council (GWAC) stack. As the issue security affects more than one interoperability layer, it is addressed as a *cross-cutting-issue* within the SGAM.

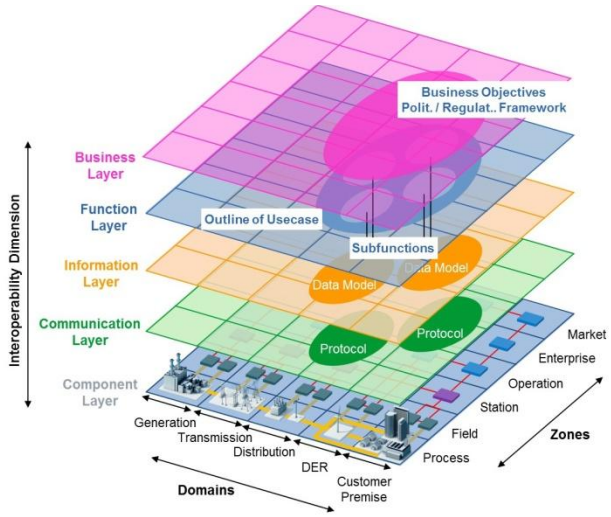

Figure 2: Smart Grid Architecture Model (SGAM)

Smart Grid Information Security

With the *Smart Grid Information Security* (SGIS) report [4], the Smart Grid Coordination Group SG-CG provides a high level guidance on how standards can be used to develop Smart Grid information security.

In this light it presents concepts and tools to help stakeholders to integrate information security into daily business. The report includes the SGAM by recommending Security Level (SL) for each SGAM Domain/Zone cell given the kind of equipment used there to manage power and its maximum potential power loss associated in a global Pan-European Electrical Grid stability scenario for a given location. These SL-recommendations are depicted in Figure 3 and range from 1 to 5, where 1 is the lowest and 5 the highest SL.

SGAM plane	Gen.	Trans.	Distr.	DER	Cust.
Market	3 - 4	3 - 4	3 - 4	2 - 3	2 - 3
Enterprise	3 - 4	3 - 4	3 - 4	2 - 3	2 - 3
Operation	3 - 4	5	3 - 4	3	2 - 3
Station	2 - 3	4	2	1 - 2	2
Field	2 - 3	3	2	1 - 2	1
Process	2 - 3	2	2	1 - 2	1

Figure 3: SGIS Security Level recommendations

So here we already have a guide for the importance of security in relation to the location of a system in design in the SGAM Model. This leads to the idea of integrating these Security Levels for risk estimation within the SGAM at design-time.

RELATED WORK

The idea of including SGIS-SL for the estimation of a risk is already addressed in [5]. Following the principle *risk = probability * impact*, it proposes the calculation of risk by the following formula [3, Eq. 8]:

$$risk = SL \cdot 2^{DOE} \cdot \sum_{i=1}^{n} API_i$$

This formula already uses the SGIS Security Level reffered to as *SL* to include the impact of an attack. The variable *DOE* is an indicator for *direct operational effects* and can have the value 1 or 0. The probability of an event is included via Attack Probability Indicators (*API*), which are given in [5] exemplarily by the three categories *hacker's motivation*, *asset reachability* and *propagation of secret*.

Building upon this, the mathematical interpretation of the given formula is analyzed in [6], so that it meets the following two requirements for the resulting risk:

- The resulting risk has an absolute reference value (independent of the number of *API* included).
- The formula can be applied on qualitative values for probability and consequence of an event.

Besides these requirements, the direct operational effects are included a bit different in the new formula in order to reduce the spread of the resulting risk. Furthermore, the new formula gives the opportunity to include a weighting factor w_i for an API and calculates the risk as follows:

$$risk = (SL + DOE) \cdot \sqrt{\frac{\sum_{i=1}^{n} w_i \cdot API_i^2}{\sum_{i=1}^{n} w_i}}$$

This formula is demonstrated in [6] on a comparison of two Smart Grid technologies that pursue the same goal. For the comparison, the exemplary categories from [5] for the API were used. But as already mentioned in [5] and again addressed in [6], these API do not cover the whole attack probability and may influence each other (a high *asset reachability* may increase the *hacker's motivation*), so now we will try to give a more detailed grouping and categorization of Attack Probability Indicators.

APPROACH

In the next step, we will check whether the total probability of a risk can be identified via regarding different *Attack Probability Indicators* and their qualitative characteristic and building the total risk by a composition of the single probabilities. Thereby, we note that the risk of a system to fail is not only given by attacks on the system, but also by failure of the system due to other reasons. Therefore, we continue speaking of *failure probabilities* and *failure vectors* instead of *attack probabilities* and *attack vectors*.

The goal is to determine different failure vectors, where a probability of occurrence can be qualitatively determined and that include any risks. In [5], the example of *hacker's motivation*, *asset reachability* and *propagation of secret* was made. Regarding the vectors *asset reachability* and *hackers motivation*, it can be argued that the *asset*

reachability directly influences *hacker's motivation*, so the given vectors can't be regarded separately and in consequence do not fit for risk estimation in the proposed way.

Another approach for the failure vectors can be derived from [7]. Regarding failure vectors where a qualitative probability of failure cannot be determined, the probability can be expressed by qualitative values. NIST [7] works with the following classifications, that can be used for our approach, too.

- Very high probability (Level 10): Probability between 96 and 100 %
- High probability (Level 8): Probability between 80 and 95 %
- Moderate probability (Level 5): Probability between 21 and 79 %
- Low probability (Level 2): Probability between 5 and 20 %
- Very low probability (Level 0): Probability between 0 and 4 %

In general, [7] gives four different types of threat sources identified: *adversarial*, *accidental*, *structural* and *environmental*. These types are even more specified, for example the thread source *adversarial* is divided into the following [7]:

- Individual
  - Outsider
  - Insider
  - Trusted Insider
  - Priviledged Insider
- Group
  - Ad hoc
  - Established
- Organization
  - Competitor
  - Supplier
  - Partner
  - Customer
- Nation-State

In regard of determining qualitative probabilities of failure, it is recommendable to regard all given subgroups separately for the example of adversarial thread sources, because the probabilities might strongly differ between these subgroups.

On the other hand, [7] names for the thread source environmental the following subgroups:

- Natural or man-made disaster
  - Fire
  - Flood/Tsunami
  - Windstorm/Tornado
  - Hurricane
  - Earthquake
  - Bombing
  - Overrun
- Unusual Natural Event (e.g., sunspots)
- Infrastructure Failure/Outage
  - Telecommunications
  - Electrical Power

Within the subgroup of *Natural or man-made disaster* the failure probabilities may differ due to the physical location of a system, but in total the probability of a *Natural or man-made disaster* at all is still very low, so this group can be summarized and treated as one failure vector.

So we can conclude that the classification of thread sources from [7] delivers a list of failure vectors that covers different aspects and gives lists of sub points for further specification. The granularity of the sub points is quite high and some sub points may regarded together due to their qualitative classification of probability.

DISCUSSION

The given formula $risk = (SL + DOE) \cdot \sqrt{\frac{\sum_{i=1}^{n} w_i \cdot API_i^2}{\sum_{i=1}^{n} w_i}}$ from [6] proposes to regard the Security Level (SL) of a system given by the SGAM cell where it is located and to multiply it by a composition of Attack Probability Indicators (API). The composition of API is not made by building the geometric or arithmetic mean, but by squaring all characteristics, summing them up and taking the square root again. This kind of composition gives a weight on strong deviations of single API, so that the total risk grows strongly when one indicator has a high probability. This is desired, because for the risk it is better when all probabilities of failure are constantly low than all probabilities slightly lower except of one high probability of failure with the same impact as the other failure vectors (expressed by the SL).

Nevertheless, there is one problem with the proposed way of risk estimation: The formula first determines all failure probabilities for one SGAM cell (with the corresponding security level), composites them to a total failure probability for the addressed SGAM cell and multiplies it with the security level afterwards. This has the effect, that included failure vectors may be counted for a low security level although their occurrence may have a higher impact or they may be counted double when multiple security level are regarded.

Example

We will explain this effect on one example: Assume we have a system located in the SGAM cell Station-Distribution. This may for example be an intelligent voltage transformer. The corresponding security level from the SGIS is 2 and the indicator for direct operational effects is 1. Then we regard the failure vector of natural disaster (Flood/Tsunami, Windstorm/Tornado, Hurricane and Earthquake). The probability for such a natural disaster to occur is very low (so we have the corresponding semi-quantitative probability level 0). But now we get to the point that a natural disaster not only causes failure at the voltage transformer of the distribution grid, but also on the whole electrical power infrastructure. So a natural disaster has an impact on the

SGAM cell Station-Transmission, too. And there the security level is 4. So when we regard the security level of the voltage transformer and calculate the risk including natural disaster, and further we then regard the transmission grid and calculate the risk here, too. Then we have the probability of a natural disaster included twice, what is mathematically not correct. On the other hand, when we avoid this problem by including the probability for natural disaster only on the cell with higher risk, but exemplarily calculate the risk for a system that only applies to one SGAM cell, we skip the failure vector of natural disease. Therefore, this option is incomplete.

Conclusion

Taken together, the given formula allows us two ways to handle with failure vectors:

1. Include each failure vector on every SGAM cell.
2. Identify directly affecting failure vectors for every SGAM cell or security level.

On the one hand, the first option leads to an overassessment of failure vectors that have impact on multiple systems simultaneously like the given example with natural disasters. But on the other hand, the second option leads to an underassessment of failure vectors that have bigger impacts expressed on another SGAM cell or security level than the one in question and therefore are left out.

RESULTS

Following the discussion, we conclude that the given formula is not satisfying. This holds for both formulas introduced in [5] and in [6] as the principle of the formulas is the same and may lead to an over- or underassessment of failure vectors. The reason for this is also addressed in [8], where one of the main findings is the unbalanced risk distribution in smart grids, ranging from higher level smart grid components from utility providers to bottom level components like concentrators or smart meter.

To avoid over- or underassessment, a holistic approach is needed. In this work, we could not show an approach that works, but we could show why it does not work and can derive the following requirements for a holistic approach:

- Some threat sources do not only affect one system, but a whole infrastructure and cannot be avoided by the design of a system.
- Even though two different thread sources might affect the same system and only that system, the impact of the thread may be different. So an overall risk estimation for a system should work on risks and not regard multiple impacts or probabilities alone before bringing them together.

Yet still we have some basic approach that can be used for risk estimation:

- The total risk is a composition of risk vectors, where [7] delivers a grouping and overview of thread sources.
- To identify the impact of a failure vector, the SGIS [4] delivers a qualitative value that can be taken into account.
- Even though an approach on single SGAM cells will not work, the SGAM itself provides a holistic viewpoint of an overall Smart Grid infrastructure and should be convenient for holistic risk estimation.

REFERENCES

[1] IEC TC 8 – Systems aspects for electrical energy supply, 2015, *IEC 62559 Use Case Methodology*, IEC International Standard.

[2] Marie van Amelsvoort, 2013, “SG-Rating - Putting values on Smart Grid architectures”, *it - Information Technology*, vol. 58, pp. 29-36.

[3] J. Bruinenberg, L. Colten, E. Darmois, J. Dorn, J. Doyle, O. Elloumi, H. Englert, R. Forbes, J. Heiles, P. Hermans, J. Kuhnert, F. J. Rumph, M. Uslar, P. Wetterwald, 2012, *Smart Grid Reference Architecture*, Technical report, CEN-CENELEC-ETSI.

[4] Smart Grid Coordination Group, 2014, *SG-CG/M490/H_ Smart Grid Information Security*, Technical report, CEN-CENELEC-ETSI.

[5] C. Neureiter, G. Eibl, D. Engel, S. Schlegel, and M. Uslar, 2016, “A concept for engineering smart grid security requirements based on SGAM models”, *Computer Science – Research and Development*, vol. 31, No. 1-2, pp. 65–71.

[6] M. van Amelsvoort, M. Uslar, 2016, “Modeling of intelligent network technologies with the Smart Grid Architecture Model for a cost-benefit analysis”, *Proceedings VDE-Kongress 2016 – Internet der Dinge*.

[7] National Institute of Standards and Technology, 2012, *Guide for Conducting Risk Assessments*, NIST Special Publication 800-30.

[8] M. Kammerstetter, L. Langer, F. Skopik, F. Kupzog, W. Kastner, 2014, “Practical Risk Assessment Using a Cumulative Smart Grid Model”, *3rd International Conference on Smart Grids and Green IT Systems (SMARTGREENS)*.