Poster: Privacy Harm Analysis: A Case Study on Smart Grids

Sourya Joyee De and Daniel Le Métayer
INRIA, Université de Lyon, France
Email: sourya-joyee.de@inria.fr, daniel.le-metayer@inria.fr

I. Introduction

Smart meters can collect highly granular energy consumption data that can expose consumers to a large number of harms, of various degrees of severity and likelihood: targeted advertising by marketers, discrimination, surveillance by the government and law-enforcement bodies, burglary or kidnapping by criminals [11]. Utility providers who have already designed a system and invested in associated equipments and technologies, face huge losses when implementations cannot be fully carried out due to oppositions from consumers and/or interventions from regulatory bodies due to risks of privacy violations. A privacy impact assessment (PIA) is vital for the early identification of potential privacy breaches and for choosing the most appropriate protection measures [8]. So a data protection impact assessment (DPIA) template for smart grids was developed by the Expert Group 2 (EG2) of the European Commission's Smart Grid Task Force [3] with feedbacks from Working Party 29 [2]. Like most other works on privacy risk assessment [1], [4], [7], [9], [10], [14], it relies on the notions of feared events, vulnerabilities and threats. The Working Party 29 [2] points out that the assessment of impacts of feared events in the template is not very clear and a list of the most relevant impacts of feared events on data subjects is desirable. To carry out a true privacy risk analysis and go beyond a traditional security analysis, it is essential to distinguish the notions of feared events and privacy harms and to establish a link between them. The Working Party 29 [2] also highlights the role of the link in this context.

Our contributions. Deriving our understanding of privacy harms from the literature on smart grids [11], [12] and privacy torts and regulations [6], [13], we provide a clear articulation between harms, feared events, privacy weaknesses and risk sources and describe their use in the analysis of smart grid systems (based on our assumptions of the system). Specifically, 1) we define the notions of harms, feared events, privacy weaknesses and risk sources; 2) we establish a relationship among these notions with the help of harm trees and 3) finally, we show that our systematic and rigorous exercise lays the foundation of an unambiguous risk assessment process. We illustrate the use of harm trees for risk assessment, deciding risks to be mitigated and privacy weaknesses to be countered in priority.

Code	Privacy weaknesses
V.1	Security vulnerabilities in Meter Data Management System
V.2	Unencrypted energy consumption data processing
V.3	Unencrypted transmission of energy consumption data from home appliance to smart meter
V.4	Non-enforcement of data minimization
V.5	No opt-outs for consumers for high volume/precision data collection
V.6	Insufficient system audit

TABLE I
PRIVACY WEAKNESSES IN A SMART GRID SYSTEM

II. Background and Definitions

We assume that a smart grid system consists of different sub-systems that store, manage and process data. The types of data used by the system are: 1) identification, contact data such as name, address, meter identifier etc.; 2) information about energy consumption and 3) information related to billing. While legal scholars mostly focus on privacy harms, technical papers talk about feared events, threats and vulnerabilities. In addition, there is often a lack of clear distinction among these concepts. So we define corresponding terms here.

Definition 1 (Risk source): A risk source is any entity (individual or organization) which may process (legally or illegally) data belonging to a data subject and whose actions may directly or indirectly, intentionally or unintentionally lead to privacy harms.

Definition 2 (Privacy weakness): A privacy weakness is a weakness in the data protection mechanisms (whether technical, organizational or legal) of a system or lack thereof.

Definition 3 (Feared Event): A feared event is an event of the system that occurs as a result of the exploitation of one or more privacy weaknesses and that may lead to privacy harms. Tabe I and Table II show some privacy weaknesses and feared events for smart grids.

Definition 4 (Privacy Harms): A privacy harm is the negative impact on a data subject, or a group of data subjects, or the society (for example, from the standpoint of physical, mental, financial well-being, reputation, dignity, freedom, acceptance in society, self-actualization, domestic life, freedom

This work has been partially funded by the French ANR-12-INSE-0013 project BIOPRIV. An extended version of this work shall appear in the proceedings of the 2016 International Workshop on Privacy Engineering (IWPE'16)

Code	Feared events	Relevant scenarios
FE.1	Excessive collection of energy consumption data	Collection of energy consumption data more frequently than billing period without consumer consent
FE.2	Use of energy consumption data for unauthorized purpose, including data inference from energy consumption data	Develop detailed consumer profiles, monitoring and restricting energy usage, inferring about a person's lifestyle or habits from his energy consumption
FE.3	Unauthorized access to energy consumption data	Service technician gets access to energy consumption data

TABLE II
FEARED EVENTS IN A SMART GRID SYSTEM

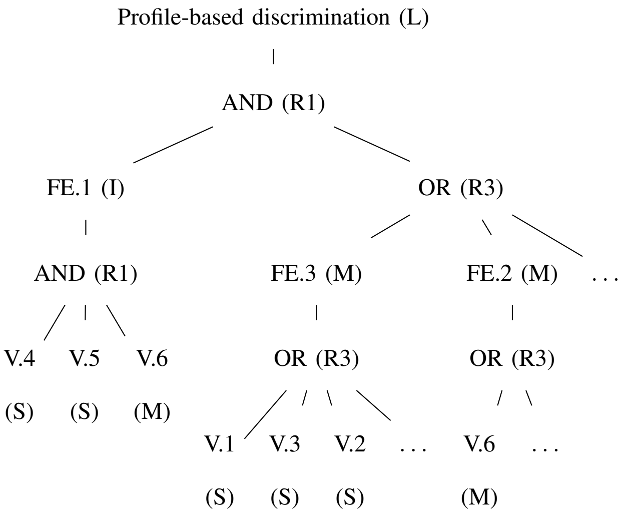

Fig. 1. Example computation of likelihood of profile-based discrimination using harm trees

of expression, or any fundamental right) resulting from one or more feared events.

Examples of relevant harms for smart grids include profile-based discrimination[1], restriction of energy usage and burglary.

III. From Privacy Weaknesses to Privacy Harms

The root node of a harm tree (akin to attack trees [5]) denotes a harm. Leaf nodes represent the exploitation of privacy weaknesses by risk sources. Feared events are connected by an AND node if all of them are necessary to lead to the harm. If any one of them may lead to a harm then they are connected by an OR node.

For risk assessment, the analyst may begin by defining the ease of exploitation of each privacy weakness for the risk sources who are most likely to exploit them. The likelihood of each harm can then be computed based on the harm tree and the likelihood of exploitation at the leaves. This process has been illustrated in Figure 1. We use the following symbolic values for input and output likelihoods (p):

1) *Negligible (N)* for $p \leq 0.01\%$;
2) *Limited (L)* for $0.01\% < p \leq 0.1\%$;
3) *Intermediate (I)* for $0.1\% < p \leq 1\%$;
4) *Significant (S)* for $1\% < p \leq 10\%$;
5) *Maximum (M)* for $p > 10\%$.

The analyst can choose other representations. The computations of likelihoods based on the harm trees rely on the following rules, where P_i is the likelihood of ith child node:

R1. AND node with independent child nodes: $\prod_i P_i$.;
R2. AND node with dependent child nodes: $Min_i(P_i)$.;
R3. OR node with independent child nodes: $1 - \prod_i (1 - P_i)$.;
R4. OR node with dependent child nodes: $\sum_i P_i$.

The severity of a harm can be obtained from the victims and the intensity of the harm. The risk level may then be represented as the pair: *(severity, likelihood)*. After the risk levels for all harms are computed, the decision maker can decide which risks are acceptable and which ones should be mitigated. A study of all harm trees for risk levels above an acceptable threshold reveal privacy weaknesses that have the strongest impact on these harms, indicating privacy weaknesses to be mitigated first.

References

[1] Privacy Impact Assessment for RFID applications. https://www.bsi.bund.de, 2011. Accessed: 2015-09-25.

[2] Working Party 29 Opinion 07/2013 on Data Protection Impact Assessment Template for Smart Grid and Metering Systems ('DPIA Template') prepared by Expert Group 2 of the Commission's Smart Grid Task Force, 2013.

[3] Data Protection Impact Assessment Template for Smart Grid and Smart Metering Systems. https://ec.europa.eu/energy/sites/ener/files/documents/2014-dpia-smart-grids-forces.pdf , 2014.

[4] Privacy risk management for federal information systems. http://csrc.nist.gov/publications/drafts/nistir-8062/nistir-8062-draft.pdf, 2015.

[5] Alessandra Bagnato, Barbara Kordy, Per Håkon Meland, and Patrick Schweitzer. Attribute decoration of attack–defense trees. *International Journal of Secure Software Engineering (IJSSE)*, 3(2):1–35, 2012.

[6] Ryan Calo. Boundaries of Privacy Harm, The. *Ind. LJ*, 86:1131, 2011.

[7] Commission Nationale de l'Informatique et des Libertes. Privacy Impact Assessment (PIA) Methodology (how to carry out a PIA), 2015.

[8] Colette Cuijpers and Bert-Jaap Koops. Smart metering and privacy in Europe: lessons from the Dutch case. In *European data protection: coming of age*, pages 269–293. Springer, 2013.

[9] Mina Deng, Kim Wuyts, Riccardo Scandariato, Bart Preneel, and Wouter Joosen. A privacy threat analysis framework: supporting the elicitation and fulfilment of privacy requirements. *Requirements Engineering*, 16(1):3–32, 2011.

[10] Jesús Friginal, Jérémie Guiochet, and Marc-Olivier Killijian. Towards a Privacy Risk Assessment Methodology for Location-Based Systems. In *Mobile and Ubiquitous Systems: Computing, Networking, and Services*, pages 748–753. Springer, 2014.

[11] Mikhail Lisovich, Deirdre K Mulligan, Stephen B Wicker, et al. Inferring personal information from demand-response systems. *Security & Privacy, IEEE*, 8(1):11–20, 2010.

[12] Stephen McLaughlin, Patrick McDaniel, and William Aiello. Protecting consumer privacy from electric load monitoring. In *Proceedings of the 18th ACM conference on Computer and communications security*, pages 87–98. ACM, 2011.

[13] Daniel J Solove. A taxonomy of privacy. *University of Pennsylvania law review*, pages 477–564, 2006.

[14] Rani Yesudas and Roger Clarke. A framework for risk analysis in smart grid. In *Critical Information Infrastructures Security*, pages 84–95. Springer, 2013.

[1] Examples include: increase/decrease in insurance premium by health insurance providers based on one's lifestyle, less favourable commercial conditions, reflection on job or loan applications etc.