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

***Abstract*—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 their impacts, called "privacy harms" here, and to establish a link between them. In this paper, we provide a clear relationship among harms, feared events, privacy weaknesses and risk sources and describe their use in the analysis of smart grid systems. This work also lays the foundation for a more systematic and rigorous approach to privacy risk assessment.**

I. Introduction

As they undertake the smart grid initiative, utility providers promise a host of benefits to consumers and the environment. The foremost advantage for consumers from installation of smart meters and roll-out of smart grids is the opportunity to reduce energy consumption enabled by detailed, personalized energy consumption reports and consequent monetary savings.

However, lurking behind the promise of better home energy management is the concern about utility providers collecting, through smart meters, highly granular (day, hour or even minutes) energy consumption data that can reveal a lot about the consumer's personal life. Research on non-intrusive appliance load monitoring (NALM) has shown that appliance level power usage data can be extracted from aggregated data collected by a smart meter based on known or learned power signatures of different appliances [25], [28], [30], [36] or even using off-the-shelf statistical tools [30]. Appliance usage profiles provide a surprisingly accurate model of different human activities [25] such as whether one cooks at home or not, sleeping patterns or when the occupants are usually away from home. These expose consumers to a large number of potentially dangerous 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 [25]. Smart meters may also provide the functionality to remotely reduce or completely disconnect energy supply [17].

Consumers and public bodies have, however, not remained silent. Protests have been directed against both the "big brother", i.e., misuse of information by the government and the "little brother", i.e., corporate misuse and commercial information resale [21]. A general survey across the US revealed that consumers had reservation about privacy of smart meters [24]. In Northern California, US, "conservatives" and "individualists" have carried out protests against installation of smart meters by the Pacific Gas and Electric Company citing various privacy concerns [10]. Initiatives such as Stop Smart Meters [7] have spread across the US, Mexico, the UK, Europe, Australia, Japan and Canada making consumers aware of various drawbacks of installing smart meters.

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. Therefore, it should be in the interest of an utility provider to carry out thorough privacy impact assessments of its system, early on, most appropriately in the design phase. A privacy impact assessment is vital for the early identification of potential privacy breaches and for choosing the most appropriate protection measures [17]. In accordance to this idea, the European Commission Recommendation 2012/148/EU issued in 2012 to provide guidance for the deployment of smart meters [3] suggests that a data protection impact assessment (DPIA) template should be adopted by Member States. This template was developed by the Expert Group 2 (EG2) of the Commission's Smart Grid Task Force [5] with feedbacks from Working Party 29 [3], [4].

Like most other works on privacy risk assessment [2], [8], [14], [15], [16], [18], [20], [44], the risk assessment methodology described by the Expert Group 2 (EG2) [5] relies on the notions of feared events, vulnerabilities and threats. The Working Party 29 [4] 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 must be provided. 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 [4] also highlights the role of this link in characterizing data protection impact assessment as opposed to an information security risk assessment. It should help understanding how and to what extent individuals (or groups of individuals, or society as a whole) can be affected by the occurrence of feared events.

A. Contributions

Although very few scholars have tried to define the notion directly [13], privacy harm is not a new concept. A large part of the privacy literature, especially those written from the point of view of privacy torts and regulations [13], [19], [29], [39],

This work has been partially funded by the French ANR-12-INSE-0013 project BIOPRIV.

DOI 10.1109/SPW.2016.21

[40] provide extensive discussions on how privacy breaches or violations or privacy harms affect the data subject as an individual or as a part of the society. Deriving our understanding of privacy harms from the literature on smart grids [25], [28], [30], [36], [43] and privacy torts and regulations [13], [27], [35], [38], we provide a clear articulation between harms[1], feared events, privacy weaknesses and risk sources and describe their use in the analysis of smart grid systems. We proceed in the following steps:

1) We provide an overview of our assumptions about the smart grid system design focusing on the energy management and billing sub-systems (Section II).
2) We define the notion of "harm" that describes the adverse impact of feared events. We also define other concepts such as "feared events", "privacy weaknesses" and "risk sources" (Section III).
3) We instantiate different attributes of harms for the smart grid scenario (Section III).
4) We then establish a clear relationship among harms, feared events and privacy weaknesses with the help of harm trees (Section IV-A) using suitable examples.
5) Finally, in Section IV-B, we show that our systematic and rigorous exercise lays the foundation to an unambiguous risk assessment process. We illustrate how harm trees can be used for risk assessment, deciding which risks need to be mitigated and selecting privacy weaknesses that need to be countered first.

Beyond its relevance for decision makers, the approach followed in this paper also enhances the accountability of the data controllers because it leads to properly documented assumptions and justifications.

II. System description

A. System design

For this case study, we assume that a smart grid system consists of the following sub-systems: 1) *User Registration System (URS)* to register new consumers with the utility provider. ; 2) *Consumer Information System (CIS)* to store and manage all consumer identification, contact and billing related information. It performs security related functions on the data stored by it. It also creates meter ID and user portal account number.; 3) *Meter Data Management System (MDMS)* to store and manage energy consumption data and corresponding meter ID. It performs security-related functions on the data stored by it.; 4) *Utility Gateway (UG)* to collect energy consumption from each smart meter. It ensures that only authorized subsystems or applications or actors can access the data collected by it.; 5) *Smart Meter (SM)* to collect energy consumption data from home appliances. It includes a security module enabling it to encrypt and sign data before sending it to the utility gateway.; 6) *Payment Management System (PMS)* to handle all billing, payment and energy management related functions.; 7) *User Interface (UI)* to enable consumers to access bills and energy management suggestions as well as update/correct any identification/contact information.

The data flows among the main components of the system are depicted in Fig. 1. The SM and the UG are located in the consumer premises. The UI can be accessed by the consumer through the Internet from his PC. All other systems are located with the utility provider and cannot be accessed by the consumer. Each new consumer registers with the utility provider using the URS by providing his identification and contact details. The URS transfers this information to the CIS which creates a meter ID and user portal account number for each new user registered. Within the consumer premises, energy consumption data from home appliances are collected at an SM. The SM then transfers this data to the UG, along with the meter ID, every 15 minutes[2] . The UG gathers data from several such smart meters. These data are then transferred to the utility provider's side to be stored and managed by the MDMS. During each billing cycle,the PMS accesses the energy consumption data for each meter ID from the MDMS and tariffs per time period given by the utility provider. The PMS computes the bill per meter ID and creates energy management suggestions based on bills and energy consumption data in each billing cycle. It also updates the payment status for each meter ID based on bills and payment information received from the bank, corresponding to a given bank account number obtained from the CIS. The resulting bill, energy management suggestions and payment status per meter ID are transferred to the CIS for storage. All data are stored and transferred in encrypted and signed form. The transfer of energy consumption data from home appliances to smart meter is, however, not secure.

B. Data processed by the system

Types of data used by the smart grid system are: 1) *Identification, contact data:* name, home address, e-mail address/phone number, date of birth, meter identifier, user portal account number; 2) *Information about energy consumption* associated with meter ID and 3) *Information related to billing:* bill and energy management suggestions (always associated with meter ID), payment status, bank account number.

For the sake of simplicity, we consider only three types of stakeholders here: the utility provider (data controller), the consumer (data subject) and the bank (third party).

III. Risk sources, privacy weaknesses, feared events and harms

While legal scholars only discuss about privacy harms, technical papers talk about feared events, threats and vulnerabilities. However, there is often a lack of clear distinction among these concepts and a clear relationship among them. In this section, we provide a definition for each of these concepts, renaming them wherever we feel necessary and describe them in the context of smart grids. In Section IV, we establish a link among them and show how the link facilitates risk assessment.

[1]We use the terms "harm" and "privacy harm" synonymously.

[2]This is the assumption made for this case study. Different choices are made in each country in which these systems are deployed.

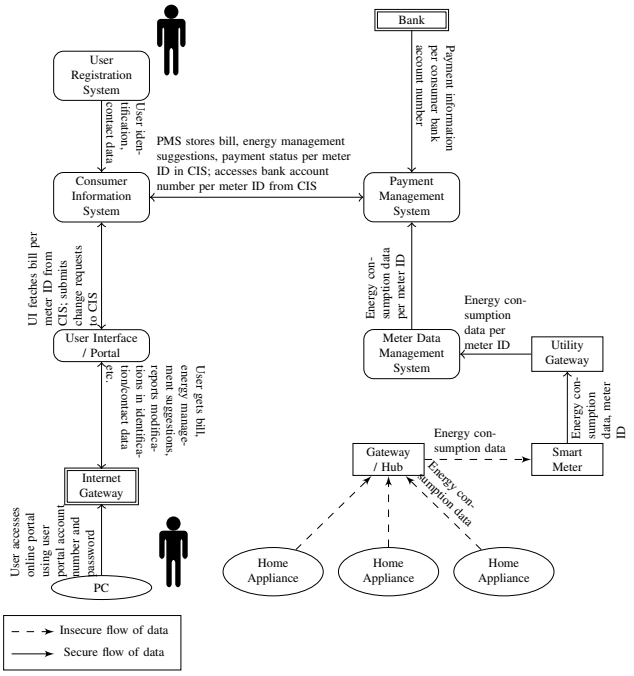

Fig. 1. Data flow diagram of a smart grid system

Code	Privacy weaknesses
V.1	Security vulnerability in PMS
V.2	Security vulnerability in MDMS
V.3	Security vulnerability in CIS
V.4	Functional errors in PMS
V.5	Functional errors in MDMS
V.6	Functional errors in CIS
V.7	Unencrypted energy consumption (per meter ID) data processing
V.8	Unencrypted billing related data processing
V.9	Unencrypted consumer identification and contact data processing
V.10	Unencrypted transmission of energy consumption data from home appliance to smart meter
V.11	Non-enforcement of data minimization
V.12	No opt-outs for consumers for high volume/precision data collection
V.13	Not assigning capabilities to consumers to challenge erroneous data about themselves
V.14	Insufficient system audit

TABLE I
PRIVACY WEAKNESSES IN A SMART GRID SYSTEM

A. Risk sources

Definition 1 (Risk source): A risk source[3] 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.

In the smart grid system, MDMS, CIS, PMS administrators, the utility provider itself or consumers, service technicians, operators or other employees, hackers can act as risk sources. Each type of risk source can be characterized by several attributes (motivation, resources, access to the system, etc.) used to assess his capacity of exploiting privacy weaknesses.

B. Privacy weaknesses

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.

Privacy weaknesses in the smart grid system can be found out from a description of existing legal, organizational and technical controls. Table I provides a non-exhaustive list of privacy weaknesses in smart grids described in the literature [11], [22], [23], [26], [31], [42]. We do not describe these in details for space considerations, since our focus is on the integration of the different notions described here in privacy risk analysis and not the discovery of privacy weaknesses. In an actual analysis process, the analyst must go into such details[4], considering both privacy weaknesses in the design (e.g. lack of encryption) and in the implementation of the system (e.g. weakness in the encryption code).

In Table I, security vulnerabilities refer to errors that lead to deviations of the system from its intended security-related functions; functional errors refer to errors that lead to deviations of the application from its intended core functions, such as wrong computation of bills; we consider unencrypted data processing separately because of the absence of such feature in the envisioned system (potential design weaknesses).

C. Feared events

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. Considering that home is the sanctuary of private life, it is clear that the installation of smart meters can lead to serious invasions of the privacy of consumers and their families [25], [28], [30], [36]. Over time, many personal characteristics such as occupation, finances, credit, health and ways of life of the occupants of a residence can be inferred from highly granular energy consumption data collected by smart meters. Unauthorized access to energy consumption or contact/identification or billing data by malicious actors exploiting privacy weaknesses in the system or disclosure to unauthorized persons can lead to gross misuse of personal data. Even utility providers themselves may engage in using the data they collect for unauthorized purposes such as selling them to third parties (e.g. data brokers). In Table II, we present a non-exhaustive

[3]They are often referred to as adversary or attacker in the literature. But, we prefer to use the term "risk sources" here as it is less security connotated and is not limited to malicious actors

[4]Any new system can be tested by experts, as usual in security, e.g. by penetration testing [41]) to find out privacy weaknesses.

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	Develop detailed consumer profiles, monitoring and restricting energy usage
FE.3	Data inference from energy consumption data	Inferring about a person's lifestyle or habits from his energy consumption
FE.4	Retaining billing related data more than required	Not deleting energy management suggestions long after consumer stops using utility provider's service, not deleting bills even after 5 years
FE.5	Retaining energy consumption data more than required	Ineffective deletion of energy consumption data from utility gateway
FE.6	Retaining contact and identification data more than required	Not deleting e-mail address, DoB even after consumer stops using utility provider's service
FE.7	Unauthorized access to identification / contact data	Hacker gets access to identification / contact data
FE.8	Unauthorized access to billing related data	One consumer gets access to another's billing data
FE.9	Unauthorized access to energy consumption data	Service technician gets access to energy consumption data
FE.10	Use of identification / contact data for unauthorized purposes	Targeted advertising

TABLE II
FEARED EVENTS IN A SMART GRID SYSTEM

list of typical feared events in smart grid systems [11], [22], [23], [26], [31], [42].

D. Harms

Definition 4 (Privacy Harms): A privacy harm is the negative impact on a data subject, or a group of data subjects, or the society as a whole, from the standpoint of physical, mental, or financial well-being or reputation, dignity, freedom, acceptance in society, self-actualization, domestic life, freedom of expression, or any fundamental right, resulting from one or more feared events.

It is important in a privacy risk analysis to choose a broad definition of harms to ensure that all possible impacts are considered, even if only specific harms prove to be relevant for a given system. The cornerstone of a privacy risk analysis is precisely the proper identification of the potential harms for the system under consideration and their severity. The useful inputs to establish this list of harms are previous privacy breaches documented or discussed in the literature (for the same type of system), case law, recommendations (e.g. published by Data Protection Authorities) and the points of view of the stakeholders. Typically, in the context of a privacy impact assessment, all stakeholders (including representatives of the subjects, e.g. civil liberty associations or privacy advocates) should be consulted and play a key role in the definition of the harms to be considered. Needless to say, some subjectivity is unavoidable in this exercise. However, all assumptions and choices should be documented and traceable. In this paper, we use as sources the harms identified in the literature on smart grids [25], [28], [30], [36], [43] and in the literature on privacy harms [13], [19], [27], [29], [35], [38], [39], [40]. Table III describes some examples of information inferred from energy consumption data of different granularities and the harms that can be caused to consumers due to such inferences. We describe some of these harms below.

The smart grid initiative has increasingly led to concerns about unwanted interference into one's private life through surveillance by the government or the law enforcement or even by other bodies [19], [29]. Even before the smart grid initiative, several instances of such surveillance have come to light. In Kyllo vs. United States [36], [39], the law enforcement used thermal imaging technique to monitor activities inside the home suspecting residential growth of marijuana. In United States vs. McIntyre [19], an investigator obtained the defendant's electricity usage records using an administrative subpoena on the suspicion of marijuana growing operations. Therefore, law enforcement can use data inferred from energy consumption data as direct or circumstantial evidences for different crimes [29]. Although this aid in criminal investigations available from energy consumption data may be considered positive as law enforcement finds it easier to perform its protective role due to advances in technology, it may still be a source of privacy violation. In other cases such as Nader vs. General Motors Corp., General Motors sought information to discredit Ralph Nader, who had charged that the former's automobiles were unsafe, through massive investigation and public surveillance [39]. With the installation of smart meters, it has become extremely easy to put anybody under surveillance without the help of any special technology or undertaking massive investigations, simply because smart meters that victims wilfully install at their homes can reveal a treasure of information.

Today, an individual's energy consumption data has become sensitive in some communities as the importance that the individual assigns to environmental responsibility [36] can be inferred from it. In 2007, the Tennessee Center for Policy Research reported that the Nashville home of the then US Vice President Al Gore, well-known as a "climate crusader", consumed significantly more electricity than the national average [36] therefore causing him embarrassment.

Any kind of release of consumer data from the utility provider could lead to a complex cascading effect for consumers and third parties, even if the latter have acquired the data legitimately, with consumer consent. In 2005, it was revealed that fraudsters posing as legitimate debt-collection firms and check-cashing companies who had used previously stolen identification and contact information to obtain valid business licenses had gained access to personal information from databases of ChoicePoint, a data broker [35].

Selling of data to third parties by utility providers is another significant concern in smart grids, as with any other data collecting system. Such third parties can be data brokers who further sell these data or information infered from them to potential employers for background checks, insurance companies, debt collection firms etc. The availability of detailed data about a consumer's daily life and habits may tempt employers to know more about potential employees even though the practice may not be ethical (or even legal in certain countries). Researchers have already noted such practices [38]

Harms	Information revealed by smart meters	Pattern	Granularity
Burglary, profile based discrimination	When are you usually away from home?	High/ low power usage during the day	Hour/ minute
Burglary	Have you been away from home for some time?	High/ low power usage during the day	Day/ hour
Kidnapping, stalking, child abuse	Do you leave a child alone at home? How often and how long?	Single person power usage or simultaneous power usage at distinct areas of the house during the day	Minute/ second
Burglary, kidnapping, stalking, profile based discrimination	Is your home protected by an electronic alarm system?	Appliance activity matching alarm system signature	Minute/ second
Profile based discrimination, Burglary	Do you own a lot of expensive gadgets?	Appliance activity matching signature of expensive gadgets	Minute/ second
Consumer profiling	Did you watch the game last night?	Appliance activity matching the game showtime	Hour/ minute
Burglary, stalking	Are you living alone at home right now?	Single person power usage or simultaneous power usage at distinct areas of the house during the day	Day/ hour
Profile based discrimination	Do you stay at home all day watching TV or in front of the computer?	Appliance activity matching signature of TV, computer	Hour/ minute
Profile based discrimination, targeted advertising	Do you cook often or prefer to eat outside?	High/ low power events around meal times for microwave, cook tops etc.	Hour/ minute

TABLE III
INFORMATION REVEALED BY SMART METERS [12], [17], [30], [36]

and resulting harms such as social anxiety in the context of social networking sites [27].

Depending on the damage caused to different aspects of the life of consumers, privacy harms of smart grids can be one of the following types:

- *Financial harms* include financial losses, damage to property etc. Burglars come to know when the occupants are not at home or if the home security system is inactive or not installed inferred from energy consumption data.
- *Psychological harms* constitute fear of misuse of personal data, fear of being observed, fear of being treated unfairly etc. A potential employer may decline a job offer to a consumer because of alleged unhealthy lifestyle inferred from his energy consumption data.
- *Harms to reputation or dignity* include embarrassment, humiliation etc. Exposure of a consumer's lifestyle may cause him embarrassment.
- *Social harms* include chilling effect or loss of creativity affecting the society due to constant monitoring by government or law enforcement bodies. Remote switching off of energy supply during periods of high demand may deprive consumers of utilities essential for leading a normal life.

A harm caused to a consumer (or a group, or society as a whole) may be a combination of any of the above types. Table IV lists some harms in a smart grid system along with their types and victims. There exists several other taxonomies of privacy harms in the context of risk assessment [6], [8], [15], [16], [32] or more generally [13], [33], [34], [39]. While some of them are based on different phases of data life-cycle and associated feared events [39], others are based on the fact that harms may be internal or external to the victim, easily measurable or not [13], tangibility, effect on society [6], or on the evolution of information tort law [34]. Although some of these works [13], [39] discuss the types of harms we refer to, the taxonomy proposed in [15], [16] which refer to physical, moral and material impacts is closest to our taxonomy.

In order to assess the severity of a given type of harm, it is useful to identify some attributes that can have a significant impact on their consequences for the victims:

a) *Victims of harms.* Different types (and numbers) of individuals can be affected by a privacy breach: 1) individual consumers or their family members (e.g., burglary); 2) specific section of consumers based on age (e.g., targeted advertising), gender (e.g., stalking of females), religion, ethnicity, profession, industry etc.; 3) society (e.g., government surveillance).

b) *Intensity.* This attribute is a composite representation of the significance of the impact on the victims. It includes the duration of the harm (from short time to irreversible), the extent of the damage, etc. A burglary can affect a consumer over a limited time due to financial losses and also lead to some psychological harms. Profile-based discrimination may have both psychological and financial effects which may last for a long time. On the other hand, targeted advertising or receipt of unsolicited mails may be irritating but do not, in general, cause any major financial or psychological harms even though they may continue over a long time.

Different measurement scales can be used to assess the severity of harms[5]. For example, the number of victims in case of profile-based discrimination can be huge, causing both psychological as well as financial distress for a long time, hence leading to a "high" severity. On the other hand, in an average scenario, burglary may lead to short term financial losses and psychological distress to some consumers and family members, hence leading to "moderate" severity. The other factor to be taken into consideration when assessing risks is the likelihood of harms, which is the subject of the next section.

IV. FROM PRIVACY WEAKNESSES TO PRIVACY HARMS

In this section, we describe our approach to establish a meaningful link among the concepts defined in the last

[5]Examples of such scales can be found in [2], [5], [14], [15], [16].

Code	Harm	Types	Victims
H.1	Kidnapping of a child	Psychological, financial	Age group
H.2	Burglary	Financial, psychological	Consumer, family
H.3	Restriction of energy usage	Psychological	Society
H.4	Profile-based discrimination	Psychological, financial	Consumers, family

TABLE IV
EXAMPLES OF HARMS IN A SMART GRID SYSTEM

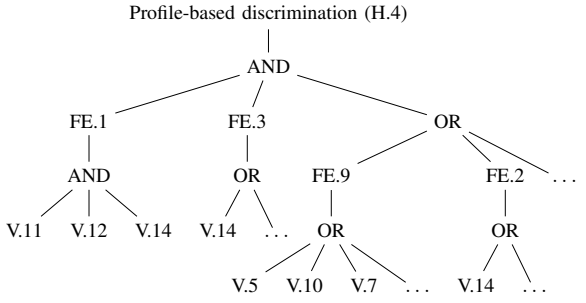

Fig. 2. Harm tree for profile-based discrimination (H.4)

section and show how this articulation can facilitate risk assessment in a smart grid system. A harm may result from one feared event or combinations of different feared events. For example, profile-based discrimination[6] can result if sufficiently fine-grained consumption data is collected, inference about personal habits, lifestyle is drawn and such data is sold to employers or others without consumer consent (or if a risk source gets access to such data). Similarly, a feared event may result from the exploitation of privacy weaknesses by risk sources. Risk sources may get access to data if data is stored, processed or transmitted without encryption or if the access control implementation is poor. One privacy weakness may lead to multiple feared events. If the utility provider does not enforce sufficient system audit, then it will be easier (because it is likely to remain undetected) to collect excessive data and to use it for unauthorized purposes. A natural way to depict such relationships among harms, feared events and privacy weaknesses is through harm trees, which are akin to attack trees in computer security literature. The use of attack trees is common in computer security [1], [9], [37] and is not new for privacy, even if very few papers have been published on this topic. In [18], Deng et al. use threat trees to link what they define as threats to vulnerabilities in a system. Similarly, Friginal et al. [20] describe attack trees to link what they define as adverse impacts (e.g., disclosure of nearest friends of an user) to attack scenarios (e.g., hacking a device). However, these works do not provide an end-to-end link between privacy harms and privacy weaknesses. For a consistent risk assessment, we assume that all privacy harms

[6]Here, we consider profile based discriminations specifically due to energy consumption data collected by smart meters. Names, postal codes etc. may be sources of discrimination, but these can be obtained in simpler ways.

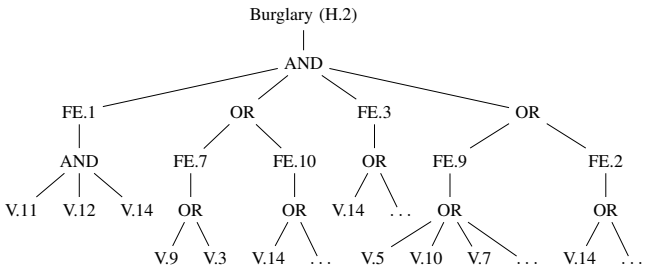

Fig. 3. Harm tree for burglary (H.2)

are caused by one or more feared events which are in turn caused by exploitation of one or more privacy weaknesses by risk sources.

A. Construction of Harm Trees

The root node of a harm tree denotes a harm. Leaf nodes represent exploitation of privacy weaknesses by risk sources[7]. The tree is structured in branches leading to the harm. Feared events are connected by an AND node if all of them are necessary to lead to the harm. For example, profile-based discrimination happens when there is excessive collection of energy consumption data, data inference and use of energy consumption data for unauthorized purposes (see Figure 2). On the other hand, if any one of several feared events lead to a harm then they are connected by an OR node. Similarly, privacy weaknesses leading to a feared event are connected by an AND node if all of them must be exploited by risk sources for the feared event to take place. Excessive data collection results when the data controller does not ensure data minimization, does not allow consumers to opt-out from such data collection and does not have sufficient system audit in place (see Figure 2). When the exploitation of any one of a set of privacy weaknesses by risk sources is sufficient to lead to a feared event then they are connected by an OR node. The observation of unencrypted energy consumption data during processing or the transmission or exploitation of a functional error in the MDM application may result in an unauthorized access to energy consumption data (see Figure 2).

Figures 2 and 3 demonstrate how harm trees should be constructed for two harms relevant to smart grid, burglary and profile-based discrimination[8], respectively.[9]

[7]Strictly speaking, a harm tree should be associated with a set of risk sources. This set can be a singleton in case of individual risk source or denote a group of risk sources, who may be colluding or not, depending on the interactions needed.

[8]Examples include: increase/decrease in insurance premium by health insurance providers based on whether one uses his treadmill everyday or eats outside frequently, by home insurance providers based on how long, how frequently one is away from home, whether one uses an electronic home alarm system etc., less favourable commercial conditions, reflection on job or loan applications etc.

[9]Dotted nodes in the trees represent the fact that there may be other feared events or other privacy weaknesses that are not pictured. We ignore dotted lines in computations based on harm trees in Section IV-B

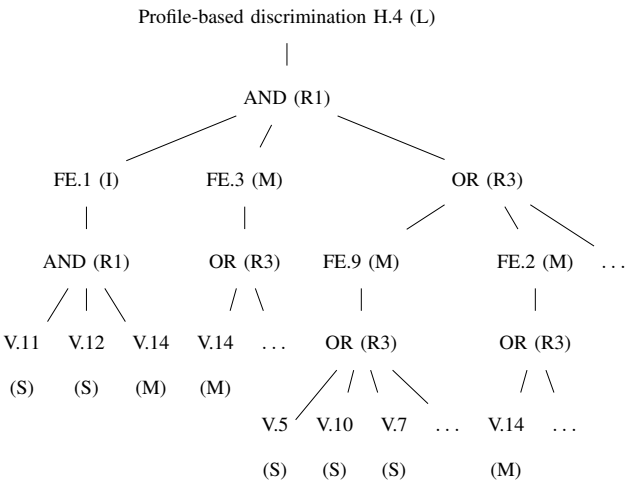

Fig. 4. Example computation of likelihood of profile-based discrimination (H.4) using harm trees

B. Risk assessment

The primary advantage of depicting the relationship among harms, feared events and privacy weaknesses through harm trees is that they can be used to compute the likelihood of different risks and eventually to decide which risks are above a tolerable threshold. The analyst may begin by defining the ease of exploitation of each privacy weakness for each type of risk source based on his attributes (incentives, expertise, resources available, etc.). The likelihood of each harm can then be computed based on the harm trees and the capabilities of the risk sources who are the most likely to interact with each privacy weakness. Likelihoods can be computed in different ways, either symbolically (based on a fixed scale of levels such as “negligible", “limited", “significant", etc.) or using numerical values (probabilities). Each approach has its benefits and drawbacks. Typically, probabilities may be difficult to estimate for input values and may look difficult to grasp by decision makers. In contrast, symbolic values are sometimes too fuzzy and may lead to different interpretations. We choose a combined approach here, with symbolic input and output values which are converted into numerical values for processing based on the harm trees and converted back into symbolic values for the final output (likelihood of the harm). We emphasize however, that the analyst can choose different representations, provided they are properly documented and justified. This process has been illustrated for discrimination in Figure 4. We use the following symbolic values for input and output likelihood (probability) values (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 computations of likelihoods based on the harm trees rely on the following rules[10], 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: $Min_i(P_i)$, i.e., minimum of the likelihoods of child nodes.; [R3.] OR node with independent but not mutually exclusive child nodes: $1-\prod_i(1-P_i)$.; [R4.] OR node with mutually exclusive child nodes: $\sum_i P_i$.; [R5.] OR node with dependent child nodes: $Max_i(P_i)$, i.e., maximum of the likelihoods of child nodes.

We have shown in Section III-D that the severity of harms can be described using two attributes, victims and intensity. The risk level of a harm may then be represented as a pair consisting of the severity and the likelihood of the harm. We observe that the likelihoods of profile-based discrimination and burglary (the computation is not described here for lack of space) are *Limited* and *Negligible* respectively. Therefore, the risk level (severity, likelihood) of discrimination is higher than that of burglary. The risk levels for other harms can be computed in the same way and the decision maker is then in a position to decide which risks are acceptable and which ones should be mitigated. The above results show that discrimination should be the primary target for risk mitigation.

A study of all harm trees corresponding to harms whose risk levels are above a given acceptable threshold also reveal privacy weaknesses that have the strongest impact on these harms. This information helps the analyst decide which privacy weaknesses should be mitigated first. Figures 2 and 3 reveal that V.11, V.12 and V.14 are necessary conditions for the occurrences of these harms. Mitigating them should therefore be a high priority. In addition, V.14 is the most commonly occurring privacy weakness, meaning that strong efforts should be put into accountability measures (especially auditing).

V. Conclusion

In this work, we have defined and instantiated important concepts in privacy risk analysis of smart grids such as harms, feared events, privacy weaknesses and risk sources and established a relationship among them through harm trees. To bridge the weaknesses in [5], we have specially emphasized on the concept of harms characterizing them through attributes derived from the literature on smart grids and privacy torts and regulations. Attributes help in determining the severity of harms. Their likelihood is derived from harm trees and the overall result of the analysis can be used to take decisions and prioritize the measures.

Beyond its relevance for decision makers, the approach put forward in this paper also enhances the accountability of the data controllers because it leads to properly documented assumptions and justifications. Indeed, even if some subjectivity is unavoidable in this endeavour (e.g. with respect to the assessment of likelihood of the feared events or the severity of the harms), all assumptions and choices should be documented and traceable. By doing so, it will be easier when the system is deployed to decide upon the appropriate corrective measures

[10]The rules are applied bottom-up to the bounds of the intervals associated with the child nodes.

in case of incident or evolution of the initial assumptions, and also to establish potential responsibilities.

REFERENCES

[1] Understanding risk through attack tree analysis. https://www.amenaza.com/downloads/docs/Methodology.pdf, 2004. Accessed: 2016-02-08.

[2] Privacy Impact Assessment for RFID applications. https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/ElekAusweise/PIA/Privacy-Impact-Assessment-Guideline-Langfassung.pdf-blob=publicationFile, 2011. Accessed: 2015-09-25.

[3] Working Party 29 Opinion 04/2013 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.

[4] 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.

[5] 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.

[6] A risk-based approach to privacy: Improving effectiveness in practice. https://www.hunton.com/files/upload/Post-Paris-Risk-Paper-June-2014.pdf, 2014. Accessed: 2015-10-01.

[7] Why stop smart meters. http://stopsmartmeters.org, 2014.

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

[9] 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.

[10] Felicity Baringer. New electricity meters stirs fear. http://www.nytimes.com/2011/01/31/science/earth/31meters.html, 2011.

[11] Kristian Beckers, Stephan Faßbender, Maritta Heisel, and Santiago Suppan. A threat analysis methodology for smart home scenarios. In *Smart Grid Security*, pages 94–124. Springer, 2014.

[12] Ian Brown. Britain's smart meter programme: A case study in privacy by design. *International Review of Law, Computers & Technology*, 28(2):172–184, 2014.

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

[14] Commission Nationale de l'Informatique et des Libertes. Methodology for Privacy Risk Management – How to Implement the Data Protection Act, 2012.

[15] Commission Nationale de l'Informatique et des Libertes. Methodology for Privacy Risk Management – How to Implement the Data Protection Act, 2015.

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

[17] 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.

[18] 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.

[19] Natasha H Duarte. Home out of Context: The Post-Riley Fourth Amendment and Law Enforcement Collection of Smart Meter Data, The. *NCL Rev.*, 93:1140, 2014.

[20] 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.

[21] David J Hess. Smart meters and public acceptance: comparative analysis and governance implications. *Health, Risk & Society*, 16(3):243–258, 2014.

[22] Marek Jawurek, Martin Johns, and Florian Kerschbaum. Plug-in privacy for smart metering billing. In *Privacy Enhancing Technologies*, pages 192–210. Springer, 2011.

[23] Nikos Komninos, Eleni Philippou, and Andreas Pitsillides. Survey in smart grid and smart home security: issues, challenges and countermeasures. *Communications Surveys & Tutorials, IEEE*, 16(4):1933–1954.

[24] David C Lineweber. Understanding residential customer support for–and opposition to–smart grid investments. *The electricity journal*, 24(8):92–100, 2011.

[25] 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.

[26] Rongxing Lu, Xiaohui Liang, Xu Li, Xiaodong Lin, and Xuemin Sherman Shen. EPPA: An efficient and privacy-preserving aggregation scheme for secure smart grid communications. *Parallel and Distributed Systems, IEEE Transactions on*, 23(9):1621–1631, 2012.

[27] Ben Marder, Adam Joinson, and Avi Shankar. Every post you make, every pic you take, I'll be watching you: Behind social spheres on Facebook. In *System Science (HICSS), 2012 45th Hawaii International Conference on*, pages 859–868. IEEE, 2012.

[28] 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.

[29] Sonia McNeil. Privacy and the modern grid. *Harvard Journal of Law & Technology*, 25, 2011.

[30] Andrés Molina-Markham, Prashant Shenoy, Kevin Fu, Emmanuel Cecchet, and David Irwin. Private memoirs of a smart meter. In *Proceedings of the 2nd ACM workshop on embedded sensing systems for energy-efficiency in building*, pages 61–66. ACM, 2010.

[31] NIST. 7628: Guidelines for smart grid cyber security. Technical report, Technical report, 2010.

[32] Marie Caroline Oetzel and Sarah Spiekermann. A systematic methodology for privacy impact assessments: a design science approach. *European Journal of Information Systems*, 23(2):126–150, 2014.

[33] Paul Ohm. Branding privacy. *Minn. L. Rev.*, 97:907, 2012.

[34] Paul Ohm. Sensitive information. *Southern California Law Review*, 88, 2015.

[35] Paul N Otto, Annie I Antón, and David L Baumer. The Choicepoint dilemma: How data brokers should handle the privacy of personal information. *IEEE Security & Privacy*, (5):15–23, 2007.

[36] Elias Leake Quinn. Smart metering and privacy: Existing laws and competing policies. *Available at SSRN 1462285*, 2009.

[37] Arpan Roy, Dong Seong Kim, and Kishor S Trivedi. Cyber security analysis using attack countermeasure trees. In *Proceedings of the Sixth Annual Workshop on Cyber Security and Information Intelligence Research*, page 28. ACM, 2010.

[38] William P Smith and Deborah L Kidder. You've been tagged!(Then again, maybe not): Employers and Facebook. *Business Horizons*, 53(5):491–499, 2010.

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

[40] Daniel J Solove. 'I've got nothing to hide' and other misunderstandings of privacy. *San Diego law review*, 44:745, 2007.

[41] Herbert H Thompson. Application penetration testing. *IEEE Security & Privacy*, (1):66–69, 2005.

[42] Wenye Wang and Zhuo Lu. Cyber security in the Smart Grid: Survey and challenges. *Computer Networks*, 57(5):1344–1371, 2013.

[43] K. T. Weaver. A perspective on how smart meters invade individual privacy. https://skyvisionsolutions.files.wordpress.com/2014/08/utility-smart-meters-invade-privacy-22-aug-2014.pdf, 2014.

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