The privacy issues for pseudonymised customers in the Smart Grid

Hartmut Richthammer[*]

Department Business Information Systems IV - IT Security Management
University of Regensburg, Germany.
Hartmut.Richthammer@ur.de

Abstract. This is a short overview of the privacy issues for customers in the Smart Grid infrastructure. Smart meter and other devices within the Smart Grid produce a lot of privacy sensitive data. With this find grained data it is possible to make predictions about the daily routine of a household or create a movement profile of a vehicle. There are techniques to protect the privacy of a person in network-alike structures, e.g. by creating pseudonyms or anonymizing the flow of data. But this protection does not always work in an adequate way, for example if there are quasi identifier, significant or linked pattern, it is possible to de-pseudonymize and de-anonymize a customer. So it is possible to analyse the daily routine of a person as well as creating a movement profile of his electrical vehicle or getting some informations about his preferences or issues.

Keywords: Smart Grid, smart meter, privacy, anonymisation, pseudonymisation

1 Introduction

A consumer, can also appear as an energy producer, as *Prosumer*. To stabilize the grid and protect it from overload and under-supply, it is necessary to keep the consumption and production of energy in an equilibrium. Therefor the Energy Service Providers (ESPs) and the Smart Meter Gateways (SMGs) of a *Prosumer* are directly connected over the Internet. As a result the acquisition of sensor data to the split second of consumed and produced energy is possible. Also controlling data can be submitted from the ESP to the *Prosumer*. So the Smart Grid initiate a paradigm change from the pure power grid to a combined power and communication grid.

This results in the following requirements for the security and privacy which are prescribed by the German BSI [7]. The Smart Grid has to be prepared to protect the ESPs against attacks. By changing the paradigm every *Prosumer* can

[*] The research leading to these results was supported by “Bavarian State Ministry of Education, Science and the Arts” as part of the FORSEC research association (http://www.bayforsec.de/).

be a victim of a distributed offense against the ESPs. On the other hand, the privacy of the *Prosumer* has to be sufficiently protected. This is necessary, because it is possible to reconstruct a detailed behaviour profile of every *Prosumer* depending on his consumption values [16,15].

2 The Privacy issue

In the future the Smart Grid will be a significant part of our life and everyone have to participate and interact with this construct. The industry wants to collect detailed, fine grained meter data of customers consumption. But each way of life and behaviour is individual. Thus a person or a household share a lot of information of its way of life with his fine granulated energy consumption trace. Figure 1 from Newborough and Augood [17] shows an example how detailed and privacy unfriendly such a trace can be. For example it is possible to determine the point in time when the user leaves the house for work and returns [15], how often the refrigerator is active and when the household have breakfast (Fig. 1) or which TV channel is watched [8].

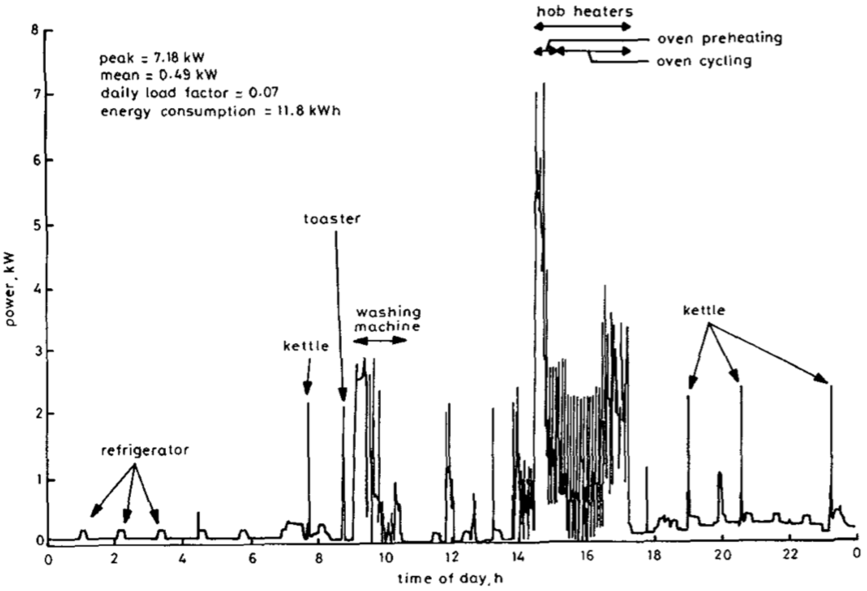

Fig. 1. Example of a Sunday electricity demand profile from an individual household. Picture cited from [17].

One additional component of the Smart Grid infrastructure will be electric vehicles because they can be integrated into the infrastructure in a useful way. The idea of the Vehicle-to-Grid (V2G) concept is to use the batteries of the vehicles as centrally coordinated, distributed grid resource [5]. But there is also a privacy problem. As Stegelmann [19] shows, a detailed movement profile can be created, because the location information of a connected vehicle is needed to manage vehicle energy flows. And also the location itself where a person parks his vehicle can reveal sensitive details. For example if it is parked frequently in front of a church, a mosque, a hospital or a medical center you can assume which belief or health condition the driver has. Predictions can be made of potential reachable destinations with the knowledge of the batteries state of charge (SOC) [19].

The privacy problem would be solved, if every consumer has the same and steady behaviour and consumption, so no individual behaviour could be identified. But this is not a realistic postulation and we have to find technical solutions, which protect the privacy on the one hand and provide the necessary data for the industry, e.g. for billing, on the other hand.

To protect the privacy of the *Prosumer*, the BSI claims that anonymisation and pseudonymisation techniques must be used [6]. Stegelmann et al. [20] describes a possible solution with the help of an example (Fig.2) as following. A Grid Operator (GO) wants to collect fine grained meter data of customers of a certain geographical area. The SMG from the *Prosumer* replaces for all non-billing relevant transmissions all identifying informations with a pseudonym. Then the data are *first* encrypted from the SMG for the GO and *then* signed by the SMG for the Gateway Operator (GWO). The GWO can now verify the signature, removes it and send the encrypted message to the GO. The GWO acts as an transportation layer anonymity service and provides assurance for the GO of the authenticity of the given data.

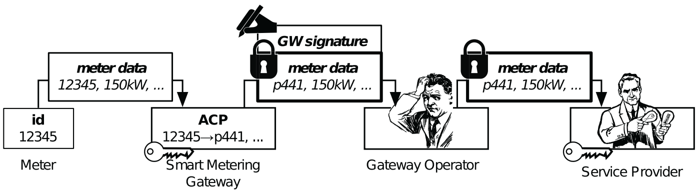

Fig. 2. Pseudonymisation and anonymisation of meter data. Picture cited from [20].

But the protection of the anonymity over a long period of time is not a trivial request and this problem is not solved yet. Because a anonymized connection does not always protect the privacy of an user. Other research work has demonstrated, that the longterm aggregation of partial information from anonymized users can break the anonymity and reconstruct the user profile [1,12,11,3]. One concrete problem and a possible solution is shown by Stegelmann et al. [20]. A customer could be de-anonymized by traffic analysis, for example based on the frequency or the absence of communication. To avoid this, enforcing information flow policies and fixed connection intervals could be used.

Another privacy protecting method is the creation of multiple pseudonyms of a *Prosumer*. The profit for the *Prosumer* would be that, if one pseudonym is uncovered, only a small part of the *Prosumer* data could be assigned to him. But multiple pseudonyms has also flaws. Jawurek [9] shows examples for possible attacks in his work. The *'linking by behaviour anomaly'* and *'linking by behaviour pattern'* attack, are shown in Figure 3.

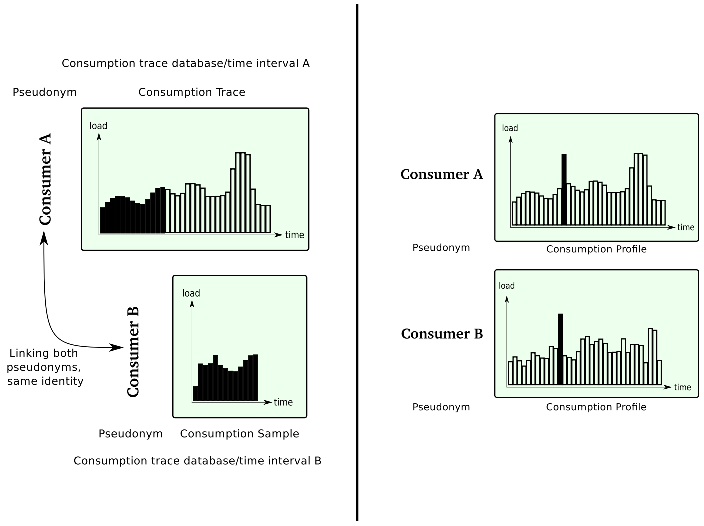

Fig. 3. Two pseudonymes are *linking by behaviour pattern* (left) and *linking by behaviour anomaly* (right). Picture cited from [9].

The *linking by behaviour anomaly* attack can be used to link either an identity to a consumption trace or more then one consumption traces together. An anomaly defines a singular or a series of unusual events. If this anomaly is reflected in the energy consumption and individual linked to the customers behaviour, an identification is possible. As an example the leaving and coming home times could be such a linking parameter. This events need to be collected in a high-resolution. But also a low-resolution identifying event could be used to identify a user. For example, if the inhabitants leave every weekend or stay at home on specific work days.

The *linking by behaviour pattern* can be used to link different pseudonyms of one person. A customer can have multiple pseudonyms. For example a new pseudonym is generated, if the supplier is changed. One other situation could be, that the supplier wants to protect the anonymity of his customer and generates a pseudonym after a certain time. For example the pseudonym A is used for the time interval t and the pseudonym B for the time interval t + 1. The benefit of this would be, that if an identity is de-pseudonymised, only a finite period of time is compromise. An attacker could now try to find a significant pattern inside the consumption traces. If this patterns are found in other consumption traces of the customers pseudonymised identities, the pseudonyms can be linked.

3 Conclusion

The Smart Grid brings an innovative change for the electricity market and his participants. But, as discussed, this change go along with risks for the privacy of every customer. It would be desirable that the *Prosumers* have the ability to protect his own privacy but for changing his individual behaviour. So the solution should be find on the modality, how the consumption data are processed and used. Jawurek et al. [10] gives a survey of different privacy technologies for Smart Grids.

4 Further Steps

The next step will be the analysis of the question, how detailed and fine granulated consumption data must be collected. One part of this step will be the investigation of Intrusion Detection Systems (IDSs). IDS are based on the analysis of (fine granulated) information flows and IDSs are necessary to protect the Smart Grid and the *Prosumer*, because there are a lot of threats [13,14,2] for example fraud and sabotage. Customer or the organized crime could try to steal energy from the ESP, a customer could try to steals energy from his neighbour or fabricate generated energy meter readings. Another scenario could be sabotage and interference, where an attacker tries to interrupt the energy supply or to destroy or damage the grid structures.

A definition of an IDS is given by [18] as a process which monitor events that occur in a computer system or network. It analyse signs of possible incidents. To detect above-mentioned threats, one solution for an IDS could be to collect and analyse consumption and behaviour data, which brings privacy issues. The question for this solution is, how detailed must this collected and analysed data be and how is it possible to combine this with privacy protection? Especially the question, is the longterm privacy protection also adequate fulfilled.

One possible method could be the anomaly detection which was early described by Denning [4]. Therefor a normal behaviour pattern from the user is established and the system looks for deviations from this behaviour. But this proceeding is not very privacy friendly and this solution always has the requirements to provide enough information to detect intruders and ensure the conservation of evidence. A decentralized intrusion detection concept, which involves the *Prosumer*, could provide more privacy. The *Prosumer* has detailed information about his own behaviour and would have an advantage in the detection of anomalies. The decentralization makes the system solid against single attacks and Single Point of Failure. As side benefit and in an ideal situation the *Prosumer* does not have to share any private data. The challenge for such a localized concept would be to detect also distributed attacks.

References

1. Dakshi Agrawal, Dogan Kesdogan, and Stefan Penz. Probabilistic treatment of mixes to hamper traffic analysis. In *Security and Privacy, 2003. Proceedings. 2003 Symposium on*, pages 16–27. IEEE, 2003.
2. Robin Berthier, William H Sanders, and Himanshu Khurana. Intrusion detection for advanced metering infrastructures: Requirements and architectural directions. In *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*, pages 350–355. IEEE, 2010.
3. George Danezis. Statistical disclosure attacks: Traffic confirmation in open environments. In *Proceedings of Security and Privacy in the Age of Uncertainty, (SEC2003*, pages 421–426. Kluwer, 2003.
4. Dorothy E Denning. An intrusion-detection model. *Software Engineering, IEEE Transactions on*, (2):222–232, 1987.
5. Xi Fang, Satyajayant Misra, Guoliang Xue, and Dejun Yang. Smart grid – the new and improved power grid: A survey. *Communications Surveys Tutorials, IEEE*, 14(4):944–980, 2012.
6. Bundesamt für Sicherheit in der Informationstechnik. Tr-03109 anforderungen an die interoperabilität der kommunikationseinheit eines intelligenten messsystems für stoff- und energiemengen, 03 2013.
7. Bundesamt für Sicherheit in der Informationstechnik (BSI). Protection profile for the gateway of a smart metering system (smart meter gateway pp), 03 2013.
8. Ulrich Greveler, Benjamin Justus, and Dennis Loehr. Multimedia content identification through smart meter power usage profiles. *Computers, Privacy and Data Protection*, 2012.
9. Marek Jawurek. *Privacy in Smart Grids.* PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg, 2013.
10. Marek Jawurek, Florian Kerschbaum, and George Danezis. Privacy technologies for smart grids-a survey of options. *Online http://research. microsoft. com/apps/pubs*, 2012.
11. Dogan Kesdogan, Dakshi Agrawal, Vinh Pham, and Dieter Rautenbach. Fundamental limits on the anonymity provided by the mix technique. In *Security and Privacy, 2006 IEEE Symposium on*, pages 14–pp. IEEE, 2006.
12. Dogan Kesdogan and Lexi Pimenidis. The hitting set attack on anonymity protocols. In *Information Hiding*, pages 326–339. Springer, 2005.
13. Zhuo Lu, Xiang Lu, Wenye Wang, and C. Wang. Review and evaluation of security threats on the communication networks in the smart grid. In *MILITARY COMMUNICATIONS CONFERENCE, 2010 - MILCOM 2010*, pages 1830–1835, October 2010.
14. P. McDaniel and S. McLaughlin. Security and privacy challenges in the smart grid. *Security Privacy, IEEE*, 7(3):75–77, May 2009.
15. Andrés Molina-Markham, Prashant Shenoy, Kevin Fu, Emmanuel Cecchet, and David Irwin. Private memoirs of a smart meter. pages 61–66, 2010.
16. Klaus J Müller. Gewinnung von verhaltensprofilen am intelligenten stromzähler. *Datenschutz und Datensicherheit-DuD*, 34(6):359–364, 2010.
17. M. Newborough and P. Augood. Demand-side management opportunities for the uk domestic sector. *IEE Proceedings - Generation, Transmission and Distribution*, 146(3):283, 1999.
18. K Scarfone and P Mell. Guide to intrusion detection and prevention systems (idps), sp-800-94. *Recommendations of the NIST National Institute of Standards and Technology (NIST)*, 2007.

19. Mark Stegelmann. *Privacy for the Smart Grid : Evaluating and enhancing Vehicle-to-Grid and Smart Metering approaches.* PhD thesis, Norwegian University of Science and Technology, Department of Telematics, 2013.
20. Mark Stegelmann and Dogan Kesdogan. Gridpriv: A smart metering architecture offering k-anonymity. In *Trust, Security and Privacy in Computing and Communications (TrustCom), 2012 IEEE 11th International Conference on*, pages 419–426. IEEE, 2012.