Embedded Cyber-Physical Anomaly Detection in Smart Meters

Massimiliano Raciti, Simin Nadjm-Tehrani

Department of Computer and Information Science, Linköping University
SE-581 83 Linköping, Sweden
[massimiliano.raciti,simin.nadjm-tehrani]@liu.se

Abstract. Smart grid security has many facets, ranging over a spectrum from resisting attacks aimed at supervisory and control systems, to end user privacy concerns while monitored by the utility enterprise. This multi-faceted problem also includes vulnerabilities that arise from deployment of local cyber-physical attacks at a smart metering location, with a potential to a) manipulate the measured energy consumption, and b) being massively deployed aiming at destabilisation. In this paper we study a smart metering device that uses a trusted platform for storage and communication of metering data, and show that despite the hard core security, there is still room for deployment of a second level of defence as an embedded real-time anomaly detector that can cover both the cyber and physical domains.

1 Introduction

Limitations of today's power networks, combined with the need for sustainable energy resources has led to promotion of smart grid architectures [1]. These promise higher reliability due to the inherently distributed nature of production and distribution, higher efficiency due to incorporation of mass scale sensors and faster management dynamics, and fine-grained adaptation to local failures and overloads. The large scale deployment of such networks is, however, dependent on exploitation of standard (IP-based) protocols, commodity sensors and actuators, and the ability of vendors to create a trusted environment on which adaptations of supply and demand can be based. The notion of cyber-physical systems, aiming to cover the "virtually global and locally physical" [2] is nowadays used even to encompass smart grids as an illustrating example.

Security is one of the less developed attributes in the cyber-physical domain. While security is indeed part of the grand challenges facing large scale development of cyber-physical systems, the focus of smart grid security is increasingly on threats to control systems [3], or serving the privacy of the end user while being subject to monitoring [4]. In this paper we address the risk of manipulations at the end-user level, even when a trusted infrastructure is assumed to be present at the smart metering end points.

The contributions of this paper are as follows:

1. We analyse the design of a smart meter which uses trusted computing technology to enforce strong security requirements, and we show the existence of a weakness in the forthcoming end-nodes, justifying real-time anomaly detection.
2. We propose an architecture for embedded anomaly detection for both the cyber and physical domains in smart meters and create an instance of a clustering-based anomaly detection algorithm in a prototype under industrial development.
3. We illustrate the detection of cyber attacks, which in principle can be script-based and massively deployed, and provide the infrastructure owner with reliable alerts.

The rest of the paper is organised as follows: Section 2 discusses the related work in this field, Section 3 presents the smart metering infrastructure, Section 4 discusses our proposed anomaly detection architecture and Section 5 shows the detection results on some cyber-attacks performed on a prototype of a smart meter.

2 Related Work

Smart grid cyber security has been a hot topic in recent years, with both researchers, industry and organisations involved in the definition of security requirements and standard solutions [5–7].

The Advanced Metering Infrastructure (AMI) is particularly vulnerable to cyber attacks, and careful attention has been given to its specific security requirements analysis [8]. Confidentiality, privacy, accountability, integrity and availability are critical requirements for accurate electricity billing and real-time power demand estimation. Cleveland [8] points out that encryption alone is not the solution that matches all the requirements, and automated diagnostics, physical and cyber intrusion detection can be means of preventing loss of availability.

Intrusion detection has been considered as a possible defence strategy in AMIs. Berthier et al. [9, 10] highlight the need for real-time monitoring in AMI systems. They propose a distributed specification-based approach to anomaly detection in order to discover and report suspicious behaviours during network or host operations. The advantage of this approach, which consists of detecting deviation from high-level models (specifications) of the system under study, is the effectiveness on checking whether the system follows the specified security policies. The main disadvantages are the high development cost and the complexity of the specifications.

Kush et al. [11] analyse the gap between conventional IDS systems and the specific requirements for smart grid systems. They find that an IDS must support legacy hardware and protocols, be scalable, standard compliant, adaptive to changes, deterministic and reliable. They evaluate a number of existing IDS approaches for SCADA systems, the approach by Berthier et al. and few conventional IDS systems that could be applied to AMIs, and they verify that none of them satisfy all the functional requirements.

Beside cyber attacks, physical attacks are also a major cause of concern. Electricity theft is the main motivation that induces unethical customers to tamper with the meters, and the minimisation of energy theft is a major reason why smart metering practice has been initiated. McLaughlin et al. [12, 13], however, show that smart meters offer even more vulnerabilities than the old electromechanical meters. Physical tampering, password extraction, eavesdropping and meter spoofing can be easily performed with commodity devices.

An approach for discovering theft detection with smart metering data is discussed in Kadurek et al. [14]. They devise two phases: during the first phase the energy balance at particular substations of the distribution systems is monitored. If the reported consumption is different from the measured one, an investigation phase aims as locating the point where the fraud is taking place.

In our environment the security requirements for AMI are fulfilled using trusted computing technology, complemented by our proposed embedded anomaly detection architecture that takes into account both the cyber and the physical domain.

3 The Trusted Smart Metering Infrastructure

The Trusted Sensor Network (TSN) [15] is a smart metering infrastructure defined as a use case within the EU FP7 SecFutur Project [16]. The main goal

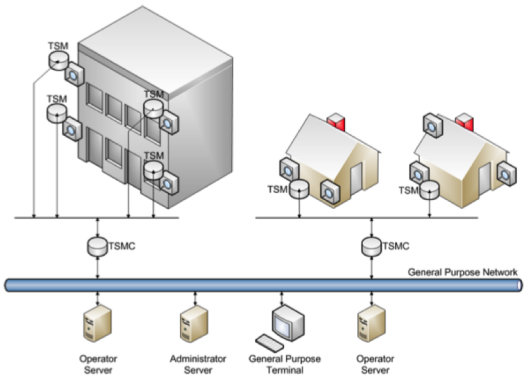


Fig. 1: Trusted Smart Metering Infrastructure [15]

of this solution is to ensure authenticity, integrity, confidentiality and accountability of the metering process in an environment where multiple organisations can operate and where legal calibration requirements must be fulfilled [15]. This goal is achieved by a careful definition of the security requirements supported by trusted computing techniques. As depicted in Figure 1, a Trusted Sensor Module (TSM) is located in each household. The energy measurements produced by one

or more sensors are encrypted and certified by the TSM, which sends them to a Trusted Sensor Module Collector (TSMC). This component gathers the data coming from several TSMs and relays it to the operator server infrastructure for its storage. Through the general purpose network several organisations can get remote access to the functionality of the metering system for installation, configuration and maintenance, but strict access policies and accountability of the actions are enforced.

A smart meter in this architecture is called Trusted Meter (TM), and it can be composed of one of more physical sensors, one or more TSMs and one TSMC. A detailed description of the architecture is presented elsewhere [15].

MixedMode^{TM}, a partner company within the SecFutur project, has developed a prototype of a Trusted Meter, described in the next section.

3.1 Trusted Meter Prototype

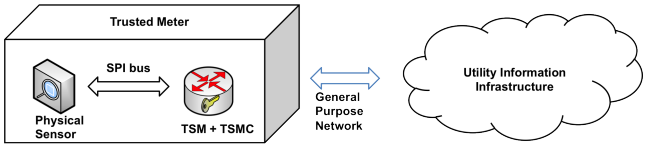


Fig. 2: Trusted Meter

The prototype of a TM is composed of one physical sensor and includes the functionalities of a TSM and a TSMC. The sensor is an ADE7758 integrated circuit, which is able to measure the accumulated active, reactive and apparent power. The functionality of the sensor is accessible via several registers that can be read or written through its interface to the Serial Peripheral Interface (SPI) bus. There is a variety of registers that can be accessed for reading out energy measurements, configuring the calibration parameters, operational states etc.

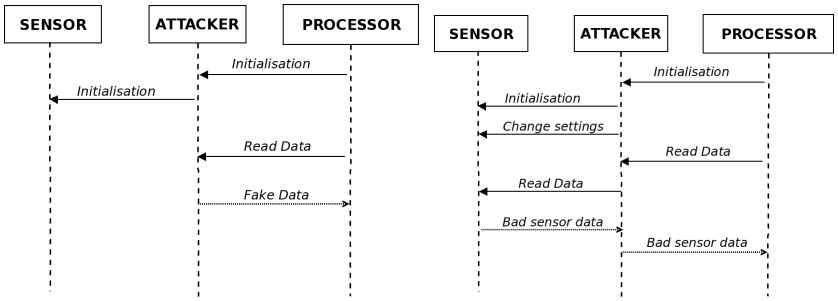
The sensor is then interfaced with an OMAP 35x system where the functionalities of the TSM and TSMC are implemented in software, with the addition of specialised hardware, namely Trusted Platform Module (TPM), that provides the trusted computing functionalities. The system, running Ångström Linux, uses secure boot to ensure that the hardware and software modules are not corrupted and encryption is used to send out the readings to the operator servers.

3.2 Threats

After a careful study of the security requirements of the system, the applied security mechanisms and the design of the meter prototype, we came up with the following observations:

– The consumption measurements, certificates, and credential recorded in the processor module (OMAP 35x) will not be subject to change by malware or external applications due to the use of TPM technology, which also prevents typical smart meter vulnerabilities reported in an earlier work [12].
– The metering data that is sent out to the operator servers is encrypted by the application, hence secure while in transmission.
– The weakest point in the system is represented by the unprotected physical connection between the sensor and the OMAP 35x system where the TSM and TSMC functionalities are implemented.

The main threat is hence represented by potential man-in-the middle attacks on the SPI bus that affect the values of data or commands transmitted, as depicted in Figure 3.



(a) Manipulation of consumption values

(b) Manipulation or injection of control commands

Fig. 3: Possible attack on the communication bus

A possible solution would be again based on encryption of the messages prior to the transmission on the SPI bus. This could be applicable in the cases when the sensor and the TSM are two physically separate modules, but when it comes to an embedded system, encryption would dramatically increase the complexity of the sensor circuitry, that must be kept cheap due to the large scale deployment.

This analysis shows the need for real-time monitoring for intrusion detection is still present although trusted platforms offer higher level of protection than earlier solutions. This motivates proposing an embedded anomaly detection as potential technology to explore.

4 Embedded Anomaly Detection

The proposed embedded anomaly detection architecture is devised to be included in the functionality of the Trusted Meter. Figure 4 illustrates the main components of the architecture. It consists of five modules: a data logger, a data preprocessor, two anomaly detection modules and an alert aggregator. The data logger

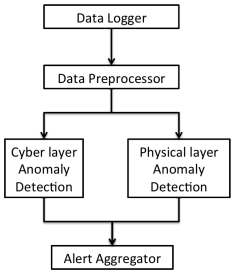


Fig. 4: Proposed Cyber-Physical Anomaly Detection Architecture

is in charge of listening for communication events and data exchange through the sensor-TSM channel. It will record both the cyber domain information, i.e. packet headers or connections, as well as the physical energy measurements.

The data preprocessor is in charge of transforming the raw signals detected on the channel into feature vectors that can be fed to the anomaly detection modules for evaluation of current state.

Anomaly detection consists of two modules: one is used for the cyber layer, i.e. the communication protocol on the SPI bus in the prototype meter, while the second is used to detect anomalies on the physical layer, the actual energy consumption that is reported by the sensor. The motivation for this distinction is the need for having two different time scales on the state estimation in the two domains: while the cyber communication can be monitored and suspicious events detected within seconds, anomalies on the physical domain need to be discovered in the order of days or weeks. This is due to the fact that load profiles change detection is only meaningful when based on a sufficiently long time window. The two modules complement each other: while command injection on the bus can be detected by the cyber layer anomaly detector, consumption data manipulation leading to changing consumption statistics can be detected by the physical layer anomaly detector.

The last component of the architecture, the alert aggregator, takes as input the alarms generated by the two anomaly detector modules and decides whether anomalous behaviour should be reported to the central system.

In the following sections we will describe the modules in depth and present their implementation in our current test system.

4.1 Data Logger

In the prototype meter, the unprotected communication channel between the sensor and the TSM+TSMC module is the SPI bus. The SPI communication is always initiated by the processor (in the OMAP 35X system) who writes, in the communication register of the sensor, a bit that specifies whether the operation is a read or a write command, followed by the address of the register that needs to be accessed. The second part of the communication is the actual data transfer from or to the addressed register of the sensor. The application that

implements the TSM and TSMC functionalities performs a energy reading cycle every second, sending calibrations or configuration commands when required.

Our bus logger records the following information: the timestamp of the operation, the command type (read or write), the register involved in the operation of the value that is read or written. In our experimental setup, as described later, the data logged at the driver level of the SPI interface of the TSM+TSMC side sufficed for our evaluation. However, bus messages should be sniffed and logged by an external element in order to record all the commands received by the ADE7758 sensor, and its deployment will be considered in the design of the next version of the prototype.

4.2 Data Preprocessing and Feature Extraction

The data preprocessor receives records in the format presented in the previous section, and produces vectors of features that will be processed by the anomaly detectors. A common data preprocessor for both domains avoids processing the received data twice. The features selected are numerical variables that all together represent the normal operation of the system. For the cyber domain, these are based on information regarding the frequency and types of operations carried out on the SPI bus during a period of observation time I. There are three categories of features:

- **Operation type**: percentage of number of read or write operations performed in the period of observation I. An additional feature counts the number of times the read-only registers are accessed, which is useful to capture the fact that most of the time (every second in our case) the communication is performed to read out energy measurements.
- **Category type**: percentage of the number of times the registers of the following categories are accessed in the period of observation I: *reading*, *configuration*, *interrupt*, *calibration*, *event*, *info*. The first category includes registers used for accumulation of active, reactive and apparent energy accumulation for the three different phases. Register categorised as *configuration* are those used for configuring different operational parameters of the energy measurement. Registers in the category *interrupt* are interrupt status flags. Registers included in the *event* category are used to store information on events such as voltage or current peak detection etc. The category *calibration*, groups the important registers used to calibrate the different parameters of the sensor. Finally, the *info* category groups registers where checksums and the version of the sensor are stored.
- **Register frequencies**: usage frequency of each individual register addressed in the period of observation I.

These features are designed to characterise the typical communication patterns, therefore anomalous communication sequences or register access rates should be discovered by the anomaly detector.

In the physical domain, commonly used indices for customers characterisation, based on load profiles, can be utilised as features. These include daily

indices, as the widely used indices proposed in Ernoult et al. [17], such as the non-uniformity coefficient $\alpha = \frac{P_{min}}{P_{max}}$, the fill-up coefficient $\beta = \frac{P_{avg}}{P_{max}}$, the modulation coefficient at peak hours $MC_{ph} = \frac{P_{avg,ph}}{P_{avg}}$ and the modulation coefficient at non-peak hours $MC_{oph} = \frac{P_{avg,oph}}{P_{avg}}$, where P_{min} is the minimum power demand reported during the day, P_{max} is the maximum power demand, P_{avg} is the average power demand, $P_{avg,ph}$ is the average power demand during the peak hours and $P_{avg,oph}$ is the power demand during the off-peak hours. More refined indices that take into account weekly patterns (working days and weekends) can be added, as those described in Chicco et al. [18].

4.3 Cyber-layer Anomaly Detection Algorithm

The sensor-processor communication is based on a series of messages exchanged through the bus. In this context, the set of possible combinations is not very large, due to the fact the set of registers accessible is bounded. The behaviour in terms of the commands sequences, captured by the features selected, can be considered as data points that fall into certain regions of the multidimensional features space. In order to identify the good behaviour, an algorithm that is able to identify these regions and consider them as the normality space would be needed.

Therefore, we have adopted and embedded an instance of a clustering-based anomaly detection algorithm [19] that uses a smart indexing strategy and is therefore computationally efficient. Section 5 presents the evaluation of this algorithm.

4.4 Physical-layer Anomaly Detection Algorithm

The features available for modelling the physical domain suggest that when the load profile changes due to an eventual attack, the statistics over a long period would be affected. A lightweight change detection algorithm can therefore be embedded into the smart meter. A statistical anomaly detector using the indicators described in Section 4.2 has been developed. However, due to absence of long term data and ability to train and test the anomaly detector on consumed electricity profiles, we have focused development and tests on the cyber level attacks.

4.5 Alert Aggregator

The last component of the architecture is in charge of collecting the alerts generated by the anomaly detector modules, and performing aggregation in order to reduce the number of alarms sent to the central operator. The alert aggregation module can gather additional information in order to provide statistics that show the evidence of an attack or the anomalous conditions. This creates a smart meter health although individual analysis would still require a lot of effort and privacy concerns would hinder its "careless" deployment. However, this

can be useful when investigating areas in which non-technical losses (i. e. losses that are not caused by transmission and distribution operations) are detected, supporting for example localisation strategies as in [14]. Future works include further investigations and privacy-aware development of this module.

5 Evaluation

In this section we present the evaluation of the anomaly detection on a number of attacks performed in the cyber domain. We start presenting the methodology to collect the data for evaluation, then we introduce the cyber attacks we performed and finally we show the outcomes of the clustering-based anomaly detection algorithm.

5.1 Data collection

In order to obtain data for training and testing the anomaly detection algorithm, the trusted meter prototype has been installed in a household and real energy consumption measurements have been collected during a period of two weeks in January 2012. Although this frame of time is not long enough to capture normality for the physical domain, it is representative enough for the cyber domain, where a 187MB data log file has been collected. The log, produced by the data logger module as explained in section 4.1, is composed of bus communication transactions that involve several registers for energy reading, sensor configurations, calibration commands and sensor events. The energy reading operations are performed with a period of one second, and they are predominant in the dataset. The data preprocessor, as presented in section 4.2, gathers the transactions during a period of observation I which has been set to 10 seconds, and produces a feature vector that is processed by the anomaly detection algorithm.

5.2 Cyber attacks and data partitioning

Four types of attack have been implemented:

1. **Data manipulation attack**: in this scenario, the attacker performs a man-in-the-middle attack in which the values of the registers involved in the energy measurement are lowered. This can be easily done by overwriting the signal on the bus on every reading cycle.
2. **Recalibration attack**: this commands is injected on the bus in order to change the value of some registers that hold calibration parameters, causing the sensor to perform erroneous measurement adjustments during its operation.
3. **Reset attack:** this command causes the content of the energy accumulation registers to be wiped out. It has to be executed within every reading cycle in order to reduce the reported energy consumption. In our scenario, we executed it with a period of one second, interleaving it with the period of the measurement process.

4. **Sleep mode attack**: this command puts the sensor into sleep mode, e.g. no measurements are taken. While in sleep mode, the sensor SPI interface still replies to the commands executed by the processor module, but the energy consumption is not accumulated by the sensor.

In our evaluation, we tested the attacks 2 to 4, since attack 1 does not produce new messages on the bus and it would only be detectable by the physical layer anomaly detector. The attacks were first implemented at application level, through the SPI interface drivers of the processor module. Since the data was collected in an attack-free scenario, a script has been implemented to weave the attack information into the clean data. Our traces consist of two weeks of logs in which 2/3 of the data represent normal conditions, and the remaining 1/3 is affected by one attack at the time, generating therefore 3 different testing traces. However, a physical hardware that can implement such attacks on the bus (SecFat) is under development in the SecFutur project.

The anomaly detector algorithm is therefore trained with the feature vectors obtained by the first third of the data, while we tested with the other 3 traces which contain the remaining third of normal data and the third of data under attack. When the data preprocessor computed the feature vectors, we manually set an oracle bit to indicate whether the features are affected by an attack or not. This will be helpful for comparison when evaluating the outcomes of the anomaly detection algorithm.

5.3 Results

During our evaluation, we have tuned the two classical parameters of the clustering-based anomaly detection algorithm which need to be configured manually in order to create a good normality model and optimise the search efficiency. These are the maximum number of clusters (M), and a cluster centroid distance threshold (E), that is used when determining whether a new data point falls within its closest cluster or not. The optimal number of clusters typically depends on the distribution of the input data into the multidimensional space. The threshold is also important, since during real-time monitoring it determines whether a new feature vector belongs to any pre-existing cluster or not. Therefore, in order to select a suitable combination of the two parameters, the outcomes of the detection were explored with M ranging from 10 to 100, and E ranging from 1 to 2.5.

The metrics used for evaluating the detection algorithm were the detection rate (DR), calculated from the percentage of feature vectors during the attack that are correctly classified as anomalous, and the false positive rate (FPR), which measures the percentage of normal observations that are erroneously classified as anomalous.

Our results show that the algorithm does not build a correct partitioning of the normality data when M is set to 10 and 20 with all the possible combinations of E. In the detection phase, all the observations (with or without attacks) are classified as anomalous, leading to 100% DR but with a 100% FPR rate. An

optimal partitioning of the normality data is found when M is set to at least 30. In this case, the algorithm uses 18 clusters to model the data, and for every configuration of E in the range between 1 and 2 we get 100% DR with no false positives for all three types of attacks. In the cases when E is over 2 we allow a very large threshold and the detection rate is reduced to zero for the recalibration attack, since it is the attack type that is more similar to normal conditions where recalibration takes place in the training period. The results are similar when increasing M up to 100. This means that the algorithm finds 18 clusters to be the best number for modelling the normality data.

6 Conclusion and future work

In this paper we have analysed the vulnerabilities of a recently designed smart metering infrastructure. Although confidentiality, authenticity, accountability, integrity and privacy are provided by the use of TPM technology embedded into smart meters, some vulnerabilities persist and real-time monitoring for cyber and physical tampering attacks is still a security solution that must be considered when designing new smart meters. Therefore, we have explored deployment of a lightweight embedded anomaly detection architecture that takes into account both cyber and physical domains and implemented and evaluated part of this architecture on a smart meter prototype. The evaluation performed on attacks in this psuedo-real settings has shown that the algorithm is able to efficiently detect several types of attacks without emitting any false positive.

Further development will target the physical layer anomaly detector and the module that combines the outcomes of the detection on both domains and provides a smart meter health indicator to provide the utility company with a more accurate non-technical loss analysis.

Acknowledgments

This work has been financially supported by the Swedish National Graduate School in Computer Science (CUGS) and the EU FP7 SecFutur Project. Support by Niklas Carstens and Maurits Broxvall, from MixedMode, collaborating in the SecFutur project, is gratefully acknowledged.

References

1. Fang, X., Misra, S., Xue, G., Yang, D.: Smart grid-the new and improved power grid: A survey. Communications Surveys Tutorials, IEEE **PP**(99) (2011) 1 –37
2. Rajkumar, R.R., Lee, I., Sha, L., Stankovic, J.: Cyber-physical systems: the next computing revolution. In: Proceedings of the 47th Design Automation Conference. DAC '10, ACM (2010) 731–736
3. Alcaraz, C., Fernandez, G., Carvajal, F.: Security aspects of SCADA and DCS environments. In Lopez, J., Setola, R., Wolthusen, S., eds.: Critical Infrastructure Protection. Volume 7130 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg (2012) 120–149

4. NISTIR 7628: Guidelines for Smart Grid Cyber Security: Vol.2, Privacy and the Smart Grid, http://csrc.nist.gov/publications/nistir/ir7628/nistir-7628_vol2.pdf, accessed Jun. 2012.
5. NISTIR 7628: Guidelines for Smart Grid Cyber Security Requirements, http://csrc.nist.gov/publications/nistir/ir7628/introduction-to-nistir-7628.pdf, accessed Jun. 2012.
6. Pallotti, E., Mangiatordi, F.: Smart grid cyber security requirements. In: Environment and Electrical Engineering (EEEIC), 2011 10th International Conference on. (2011) 1 –4
7. Lu, Z., Lu, X., Wang, W., Wang, C.: Review and evaluation of security threats on the communication networks in the smart grid. In: Military Communication Conference, 2010 - MILCOM 2010. (2010) 1830 –1835
8. Cleveland, F.: Cyber security issues for advanced metering infrastructure (ami). In: Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, 2008 IEEE. (July 2008) 1 –5
9. Berthier, R., Sanders, W., Khurana, H.: Intrusion detection for advanced metering infrastructures: Requirements and architectural directions. In: Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on. (Oct. 2010) 350 –355
10. Berthier, R., Sanders, W.: Specification-based intrusion detection for advanced metering infrastructures. In: Dependable Computing (PRDC), 2011 IEEE 17th Pacific Rim International Symposium on. (Dec. 2011) 184 –193
11. Kush, N., Foo, E., Ahmed, E., Ahmed, I., Clark, A.: Gap analysis of intrusion detection in smart grids. In Valli, C., ed.: 2nd International Cyber Resilience Conference, secau - Security Research Centre (August 2011) 38–46
12. McLaughlin, S., Podkuiko, D., McDaniel, P.: Energy theft in the advanced metering infrastructure. In: Proceedings of the 4th international conference on Critical information infrastructures security. CRITIS'09, Springer-Verlag (2010) 176–187
13. McLaughlin, S., Podkuiko, D., Miadzvezhanka, S., Delozier, A., McDaniel, P.: Multi-vendor penetration testing in the advanced metering infrastructure. In: Proceedings of the 26th Annual Computer Security Applications Conference. ACSAC '10, ACM (2010) 107–116
14. Kadurek, P., Blom, J., Cobben, J., Kling, W.: Theft detection and smart metering practices and expectations in the netherlands. In: Innovative Smart Grid Technologies Conference Europe (ISGT Europe), 2010 IEEE PES. (2010) 1–6
15. Broxvall, M.: Metering devices with legal calibration requirements, Deliverable D2.1, http://www.secfutur.eu/WP2/Deliverables/D2.1/Deliverable_D2_1.pdf, accessed May 2012.
16. EU SecFutur FP7 Project: http://www.secfutur.eu, accessed Jun. 2012.
17. Ernoult, M., Meslier, F.: Analysis and forecast of electrical energy demand. In: Revue générale de l'électricité. Volume 4. (1982)
18. Chicco, G., Napoli, R., Postolache, P., Scutariu, M., Toader, C.: Electric energy customer characterisation for developing dedicated market strategies. In: Power Tech Proceedings, 2001 IEEE Porto. Volume 1. (2001) 6 pp. vol.1
19. Burbeck, K., Nadjm-Tehrani, S.: Adaptive real-time anomaly detection with incremental clustering. Information Security Technical Report - Elsevier **12**(1) (2007) 56–67