Privacy Preserving Energy Management

Holger Kinkelin[1], Marcel von Maltitz[1], Benedikt Peter[1], Cornelia Kappler[2],
Heiko Niedermayer[1], and Georg Carle[1]

[1] Technische Universität München, Boltzmannstr. 3, 85748 Garching, Germany
[lastname]@net.in.tum.de
[2] deZem GmbH, Sybelstr. 63, 10629 Berlin, Germany
[firstname.lastname]@dezem.de

Abstract. The improvement of energy efficiency is an important target on all levels of society. It is best achieved on the basis of locally and temporally fine-grained measurement data for identifying unnecessary use of energy. However, at the same time such fine-grained measurements allow deriving information about the persons using the energy. In this paper we describe our work towards a privacy preserving system for energy management. Our solution follows the privacy by design paradigm and uses attribute-based cryptography and virtualization to increase security.

1 Introduction

The increase of energy efficiency is an important target on all levels of society. According to studies [SZ11] [V. 09], it is best achieved when users receive information about their energy consumption such as how much is consumed, when, by which device, in what form. Furthermore, the effectiveness of this measure is best, when information is given close to the point in time energy is consumed [Fis07]. The smart meters increasingly installed in European homes, reporting consumption with a resolution of 15 minutes, are a first step in this direction.

What is true of the individual energy user is also true for *energy managers*, responsible for decreasing the energy consumption of their organisation, e.g. a company, an office building or a factory: In this context, the collection of energy consumption measurements is usually embedded in an *energy management system* (EMS) such as ISO 50001 [fSI11], which provides a continuous improvement process supporting the discovery and realization of energy saving potentials. Consumption measurements are compiled using a wide variety of methods, ranging from manual reading of meters to data loggers automatically providing measurement data collected from buildings equipped with thousands of sensors.

A central parameters in the design of an EMS is the resolution of the measurements, both in space and in time. The low end of the spectrum is the single aggregated consumption number manually collected once per year. The advanced end of the spectrum provides real-time measurement data with a temporal resolution of seconds and spacial resolution down to the individual device, plus additional information e.g. on weather or building and device status. The better the resolution, the more targeted and efficient the energy saving measures. For example, a heating system which is not configured properly and thus working at inappropriate hours can be discovered quickly, just as open fridge doors or lights left on.

The problem addressed in this paper is that an EMS can be abused and turned it into a system that monitors e.g. employees via their energy consumption. In Fig. 1 we depict the power consumption of a computer workplace, which shows that it is easy to derive information about the user's workday.

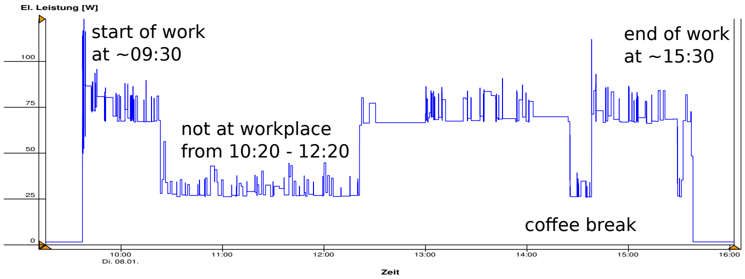

Fig. 1. Detailed energy monitoring tells much about a user's behavior.

Correlating energy logs with other information sources, e.g., the system keeping track of working times of employees, is just a further step. It is problematic that the time an employee spends working at her computer can thus be compared to her "claimed" working time.

We conclude that the high resolution measurements (desirable and necessary for improving energy efficiency) and personal privacy appear to be contradicting requirements. When an EMS based on high-resolution measurements is established in an organization, usually the work council is involved and any objections are addressed. This makes the introduction of the EMS more complicated and eventually may endanger its successful operation.

In this paper we present our work towards an EMS that tries to resolve this apparent contradiction. The solution can be applied to similar use-cases where access rights to privacy-sensitive data streams have to be enforced. We first detail the problem in Chapter 2 and outline background information in Chapter 3. Our approach is discussed in chapters 4 and 5.

2 Analysis

2.1 Are Energy Logs Personal Data?

According to the European Data Protection Directive, *personal data* is defined as "any information relating to an identified or identifiable natural person" [Par95]. Especially when office rooms are used by only one person it is straight forward to relate energy logs of this room to this person. Hence, energy logs need to be considered as personal data and data protection laws have to be applied.

2.2 EMS User Groups and their Access Rights

Before introducing different roles in the EMS, we must introduce several assumptions about a building. A building is provided and managed by an authority we call *Building Supplier* (BS). Different *spaces* of the building (multiple floors, rooms, etc.) can be rented by *Building Customers* (BC), for instance, companies. BCs pay money to the BS for rented space and consumed energy (electrical power, heating, etc.).

Based on above assumptions, we introduce user groups in the EMS. Each user is represented in the EMS by processes that automatically analyze data accessible for this user.
Building User (BU): a person working in the building. By default she is allowed to retrieve energy measurement data in full granularity concerning herself. A BU belongs to a BC. *Energy Manager* (EM): belongs to a BC. Her duty is finding unnecessary energy consumption in the company and optimizing the situation. Hence she needs to access detailed energy logs of the entire space rented by the BC. This data set comprises personal data of several BUs. *Energy Accountant* (EA): belongs to the BS. Her tasks include billing of energy by BCs. For the EA a spatially and temporally aggregated view on energy logs, i.e., the sum of energy spent within the space of an BC, is sufficient.

2.3 Design Strategies for Privacy by Design vs. Energy Control

Hoepman defined in [Hoe12] eight design strategies for privacy by design systems. These strategies are derived from data protection laws, such as [Par95]. In the following we introduce the most relevant strategies and assess their applicability.

Strategy #1 - Minimize: "The amount of personal data that is processed should be restricted to the minimal amount possible." [Hoe12] As already described, collecting data with high resolution is essential for an EMS. Otherwise, the successful operation of the EMS is compromised. Hence, this strategy can not be applied.

Strategy #2 - Hide: "Any personal data, and their interrelationships, should be hidden from plain view." [Hoe12] This design strategy does not conflict with an EMS and can be implemented by access control mechanisms.

Strategy #3 - Separate: "Personal data should be processed in a distributed fashion, in separate compartments whenever possible." [Hoe12] This strategy does not conflict with an EMS. However, distributed processing and storage has a bigger overhead in administration than a centralized system.

Strategy #4 - Aggregate: "Personal data should be processed at the highest level of aggregation and with the least possible detail in which it is (still) useful." [Hoe12] Aggregation will reduce resolution. Hence, this strategy is basically equal to strategy #1 and can only be applied in selected cases, e.g. for the EA.

Furthermore a system following privacy by design should also obey the following remaining principles. It should **inform** (#5) a user when her personal data is processed as well as preferably putting her in **control** (#6) of her data and its distribution. Furthermore, the system's compliance to legal requirements must be **enforced** (#7) while being able to **demonstrate** (#8) this property.

3 Background

3.1 (Hardware) Virtualization

Hardware virtualization is a technology able to execute several *virtual machines* (VMs) on the same physical hardware, e.g., a server. An individual operating system and any desired set of applications can be installed in a VM. As a result, virtualization is a quite popular technology today and used to cut down hardware costs in companies, or as basis for Cloud computing. [Cit12]

Besides the direct benefits of virtualization, it is often used as a security mechanism when several critical processes on the same machine need to be isolated. Instead of only relying on the operating system's ability of processes isolation, the virtualization system adds another isolation layer. This layer makes it more difficult to take control over one process after compromizing the other. However, isolation by virtualization is not impeccable. [RW10]

3.2 Attribute Based Encryption

Attribute based encryption (ABE) is a crypto system initially proposed by Sahai and Waters [SW04]. It consists of a trusted Key Generator (KG), which initially creates and owns a master key and a public key. The KG's purpose is to create private keys for other users using its master key. These private keys will include attributes of its owner, e.g., her identity, her security clearance, etc.

Besides their private key, users of an ABE crypto system possess the global public key, which allows them to encrypt data. A unique property of a specific ABE type, called *cipher text policy ABE* (CP-ABE) [BSW07], is that a policy is integrated into the cipher text that expresses who is able to decrypt it. Policies can include required attributes (attribute `admin` is set), inequalities (`clearance > 3`), and even complex boolean expressions. An entity that tries to decrypt data will only succeed if the attributes of her private key conform to the specified policy. Hence, CP-ABE offers powerful, cryptography-based access control to data.

4 Approach

4.1 Abstract EMS Architecture

A high-level architecture showing the most important components of an EMS is depicted in Fig. 2. A *data logger* equipped with *sensors* measures electrical consumption. It outputs a stream of data elements, each consisting of the logger's identity, the identity of the sensor, the measurement value and a time stamp.

Next, the data stream is pre-processed and stored. The *combiner* enriches the data stream with additional *knowledge* about the monitored system. For instance, the combiner knows that sensor "a" attached to data logger "x" measures the electrical consumption in BU Dave's office. Hence, the combiner assigns the measurement value to Dave. The enriched data stream is finally stored within a data sink from where it can be accessed after authorization by the different user groups defined before.

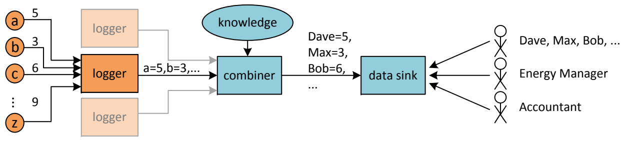

Fig. 2. Abstract EMS Architecture.

4.2 Limitations of Privacy Enhancing Services

Typically, privacy protection would be added by a separate layer or module implementing e.g. access control mechanisms. This approach has benefits when existing systems need to be extended with privacy preserving features. The downside is, that circumvention of the mentioned module allows raw access to the unprotected data (e.g. the database files on the server). In consequence the effectiveness of privacy preservation does not only depend on the given mechanism but also on the security of the overall server infrastructure.

4.3 Logically Distributed EMS based on CP-ABE and Virtualization

Our approach to create an EMS architecture with privacy by design is based on the idea to combine CP-ABE with isolation offered by hardware virtualization. Above architecture can simply be extended after the combiner, see Fig. 3.

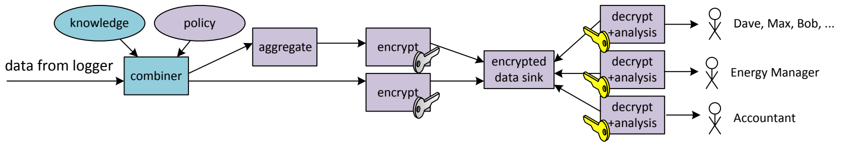

Fig. 3. Privacy by Design EMS Architecture.

In order to allow processes acting for a specific user to find accessible data elements, CP-ABE encrypted index structures are created for every user. A pointer as well as a random AES key are created and appended to this index document. In the cipher text, a policy is embedded which makes sure that only a specific BU, e.g., Dave, and (EM Dave), the EM of the company Dave belongs to will be able to decrypt the data element. In the example Dave and (EM Dave) would both get their own index document containing the same key and a pointer.

After the combiner assigned a data element to a BU, the *encrypter* component encrypts the data using AES and the AES key. The encrypted data element is then stored within a database using the pointers as (unencrypted) index fields. By sharing the AES key, each data point has only to be stored once, even if multiple users are allowed to retrieve it. Other incoming data elements are processed analogously.

While a process with a valid private key is able to find encrypted data blocks, it is not possible to do reverse queries. Therefore, one cannot gain information on who may decrypt a specific data block. This is an important feature, because otherwise it would be possible to link appearance of persons to specific points in time (e.g. Dave and Bob were present in the same room).

Here the strategy of data *hiding* is applied: The energy data itself is protected from unauthorized access as well as the access policy.

Hardware virtualization is used to isolate processes concerned with above processing of incoming data from processes that analyze data. For every user of this system an individual VM is provided. Within a VM, the private CP-ABE key of the user resides, which allows the analysis processes to decrypt data elements accessible by this user. Please note, data analysis might also be performed on a different physical machine. This allows the desired *separation* of critical data.

Aggregation of data elements, e.g., to provide data needed by the EA, is also performed within a separate VM. After assigning an incoming data element to a BU, the combiner sends the unencrypted data element to the responsible aggregation process. This process understands that the element belongs to a BC it is responsible for and adds the new value to the sum of energy spent by all BUs that belong to the BC. The element itself is then discarded. After the applicable aggregation period is over, the aggregation process encrypts the aggregated value for the EA and stores the result in the database. This part realizes the *aggregation* of data to remove undesired fine-granularity of information. This approach also supports, to a small degree, the *minimization* of the amount of persisted data.

5 Evaluation and Discussion

As we are describing work in progress we cannot provide a thorough performance analysis yet. However, our first tests indicate that our reference implementation running on a machine with an Intel Core 2 Duo CPU with 2.5Ghz, 8GB of RAM and a conventional hard drive easily copes with the data stream produced by the energy logger deployed at our group. Additional tests showed that we are able to scale it up to 2 Million data elements per hour without performance issues.

The increased complexity of our EMS (privacy by design, CP-ABE-based access control; logically distributed) compared to a simple EMS (privacy functions, access control by authentication and authorization; centralized) is worth the effort when we compare security features of both systems. In case an attacker gains access to the simple EMS, she would be able to access the database containing energy logs of the entire building and could eavesdrop on incoming data elements. In our EMS we need to differentiate. In case the attacker gains access to the VM of a BU, only this BU's data elements are accessible. In case the attacker gained access to an EM's VM, all data elements accessible by this EM can be accessed. The logger's data stream is not affected in both cases. In case a VM with an aggregation process is compromised, the attacker will be able to eavesdrop all data elements sent to the aggregation process. At the worst, this would be the energy consumption of the entire building. However, the attacker is unable to access detailed energy logs of the past. Compromising the VM running the combiner and encryptor would have a similar effect on incoming data elements. However, the attacker has no access to past energy logs.

Our system is protecting measurement data and not meta data of requests. An observer logging user requests will not see the data, yet she can observe if a user (in the sense of IP address) is requesting data for a given time interval.

As explained, our EMS is able to effectively limit the effects of attacks and increase the effort an attacker must spend to gain access on all data. For increased security, VMs can be monitored by Host Intrusion Detection Systems (HIDS). When the HIDS detects that a VM is compromized, the VM can be shut down in order to prevent all unencrypted information from being compromised. An identical VM can then be restarted from an integrity protected VM image.

6 Related Work

Smart meters report the power consumption of a building e.g. every 15 minutes to the energy provider. However, even this information is sufficient to allow the energy provider an overall footprinting of the household. To circumvent this problem, technologies such as homomorphic encryption are used. The basic idea of such systems is that multiple parties, e.g., the smart meters located in one street, cooperate in order to calculate the sum of consumed power within this street. For the energy provider this information is still sufficient in order to control energy production. Secure multiparty computation using homomorphic encryption [GJ11] now guarantees that neither the individual smart meters nor the energy provider learn about the energy consumption of single households, as all computations are performed on encrypted values. This type of work is an example where privacy can be increased by aggregation.

7 Conclusion and Future Work

Energy management systems (EMS) are needed in smart buildings to decrease energy consumption. However, there is a conflict between the need for fine-grained energy measurements and the right on personal privacy of affected persons. We analyzed strategies for designing privacy preservation in systems processing personal data and assessed their applicability on EMS. We then outlined our approach which combines the features of attribute-based encryption to enforce strong access control and of virtualization to gain isolation of processes handling personal data. A first evaluation of our prototype's performance showed that it is easily capable to handle the energy measurements of the installation in our building. Furthermore, a discussion of security features showed the clear advantages of our system compared to a centralized system.

Up to now we addressed mechanisms that increase privacy properties of an EMS. However, to achieve compliance to privacy regulations further properties, such as informing a user and demonstrating the system's abidance to legal requirements need to be addressed, i.e. to have no secret surveillance functions. We plan to address this issue based on our preliminary work on Trusted Computing and attestation technologies [KDC14].

Acknowledgments

This work has been supported by the German Federal Ministry of Education and Research, projects IDEM (grant 01LY1217C) and BaaS (grant 01IS13019G) as well as by the European Commission, FP7 project EINS (grant 288021).

References

[BSW07] John Bethencourt, Amit Sahai, and Brent Waters. Ciphertext-policy attribute-based encryption. In proc. of the 2007 IEEE Symposium on Security and Privacy, 2007.

[Cit12] Citrix Systems, Inc. The Xen Project. Website, 2012. Available online at http://www.xen.org/; last accessed on 2013/07/09.

[Fis07] Corinna Fischer. Influencing Electricity Consumption via Consumer Feedback. In proc. of the ECEEE 2007 Summer Study, June 2007. Available online at http://www.tips-project.org/download/TIPS_DP8_Fischer.pdf; last accessed on 2014/07/08.

[fSI11] International Organization for Standardization (ISO). Energy management systems requirements with guidance for use (iso 50001:2011);, June 2011. Available at http://www.iso.org/iso/home/standards/management-standards/iso50001.htm; last accessed on 2014/09/16.

[GJ11] Flavio D Garcia and Bart Jacobs. Privacy-friendly energy-metering via homomorphic encryption. In *Security and Trust Management*, pages 226–238. Springer, 2011.

[Hoe12] Jaap-Henk Hoepman. Privacy design strategies, 2012. Available online at http://arxiv.org/abs/1210.6621; last accessed on 2014/07/08.

[KDC14] Holger Kinkelin, Michael Dorner, and Georg Carle. Lokale Integritätsverifikation von Systemen durch Java Smart Cards. Tagungsband des 24. SmartCard Workshop, February 2014.

[Par95] The European Parliament. Directive 95/46/EC. Website, November 1995. Available online at http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML; last accessed on 2014/07/08.

[RW10] Joanna Rutkowska and Rafal Wojtczuk. Qubes OS Architecture Version 0.3. White paper, Invisible Things Lab, January 2010.

[SW04] Amit Sahai and Brent Waters. Fuzzy identity based encryption. Cryptology ePrint Archive, Report 2004/086, 2004. Available online at http://eprint.iacr.org/2004/086; last accessed on 2014/07/08.

[SZ11] Amt für Hochbauten Stadt Zürich. Schlussbericht Nutzerverhalten beim Wohnen. Website, 2011. Available online at http://www.mehralswohnen.ch/fileadmin/download/1107_Bericht_Nutzerverhalten.pdf; last accessed on 2014/09/18.

[V. 09] V. Bürger, Öko-Institut. Identifikation, Quantifizierung und Systematisierung technischer und verhaltensbedingter Stromeinsparpotenziale privater Haushalte. Study, 2009.