TOWARDS PRIVACY IN MONITORED SHARED ENVIRONMENTS

Kaibin Bao[1], Thomas Bräuchle[2], and Hartmut Schmeck[1]

{bao, braeuchle, schmeck}@kit.edu
Karlsruhe Institute of Technology (KIT),
[1] Institute of Applied Informatics and Formal Description Methods (AIFB),
Kaiserstr. 89, 76133 Karlsruhe (Germany)
[2] Center for Applied Legal Studies (ZAR),
Vincenz-Prießnitz-Str. 3, 76131 Karlsruhe (Germany)

Keywords: privacy, visualization, monitoring, energy data, access rights, surveillance.

1 PRIVACY THREATS OF VISUALIZING ENERGY AND SENSOR DATA

For the purpose of energy efficiency, automation and consumption visualization, an increasing number of sensors like power meters, temperature sensors, and CO_2 sensors, are installed in private households. Such a monitored household, the Energy Smart Home Lab (ESHL)[1], was built on the campus of KIT for research on energy management systems [1]. In addition to ambient temperature and thermal flows, the lab records energy consumption data of 37 electric metering points to monitor each appliance. The recorded data can be visualized in the form of graphs as shown in Fig. 1. Monitored environments like the ESHL are usually shared by multiple residents whose actions and behavior are reflected in the data. By recording and analyzing the data, inferences about the actions and habits of residents of a household can be drawn. Molina-Markham [2] and Lisovich [3], for instance, demonstrated that sensitive information like occupancy, appliance usage, sleeping cycles, and cooking habits, can be inferred using only energy consumption data. On the one hand, revealing such sensitive information to the involved residents is very important to improve energy awareness. On the other hand, sensitive information about other residents can be abused for continuous surveillance.

This work presents a novel concept to defining access rights within an energy monitoring and visualization system capable of actively mitigating the ability for mutual surveillance.

1.1 Scenario and assumptions

In order to define access rights on energy and sensor data, we assume the following use case and application context:

1. An energy and sensor monitoring and visualization system is installed in an environment which is shared by multiple residents. The monitoring aims at finding out how the energy efficiency of their building can be improved, maybe by replacing or repairing appliances, or pointing out wasteful behavior of the residents or of the other activities or processes in the building.
2. The monitoring system consists of multiple sensors, a database to store the energy and sensor data as time series and a visualization frontend, which generates different views on the stored data.
3. The residents are most importantly interested in their own energy footprint as well as long-term trends in total energy consumption of the whole household and of each individual appliance.

[1] http://www.izeus.kit.edu/english/57.php

1.2 Threats and attacks on the privacy of individual residents

Access control in current energy monitoring and visualization systems like Discovery[2], Plotwatt[3], or the web interface of the ESHL, is usually defined on the granularity of whole time series only. In most cases, access rights are granted by the all-or-nothing-principle on complete households or buildings. Such systems allow all residents to observe the complete recorded data in highest resolution. This data can be used to easily determine the time periods when appliances are being used, effectively seeing the actions happening inside the whole household or building. Therefore, all residents having access to the visualization system can abuse the system for surveillance. This kind of continuous surveillance is ubiquitous within a building and is hardly noticeable by human senses. Hence, it poses a serious threat to privacy. As the visualization providers offer access over the internet, even remote surveillance is possible. Thereby, the intensity of surveillance may exceed the level of socially acceptable observation (e.g. an employer supervising his/her employee, parents keeping an eye on their children) and is comparable to constant video surveillance.

From a legal perspective, a legitimation by law or the consent of the affected person to be monitored is required, due to the right to informational self-determination. In order to determine the own behavior free of any suppression, people have the right to decide who should receive which personal information. If people are not able to estimate the knowledge of her/his social environment or potential communication partners concerning their personal detail with sufficient certainty, they can be substantially inhibited in their freedom to make plans or decisions out of their own determination and live according to those plans and decisions. [4]

2 REFINING ACCESS CONTROL TO IMPROVE PRIVACY

The key idea to prevent surveillance is to link the right to access energy and sensor data with the presence of a particular resident. Data recorded during a specific time can only be accessed by residents whose presence was detected at that time.

The access rights can be further refined by dividing the monitored environment into independent domains. Domains can be defined by spatial or ownership properties, for example, if each children should not be able to see what is going on in the other children's room. Domains are also needed to separate office rooms.

We argue that access to highly aggregated data does not violate the privacy of individuals. Also, highly aggregated data provides important information like long-term trends in total energy consumption. The visualization system should display highly aggregated data as a fallback whenever fine-grained data is not accessible.

2.1 Requirements

The proposed privacy enhancement can be summed up to the following requirements:

1. Access to fine-grained energy and sensor data of a domain is only granted if presence of the resident was recorded during the time of recording.
2. The monitoring system thus has to keep track of the presence of particular residents in each domain.
3. Access to energy and sensor data for all residents is only granted as aggregated values. This is needed to calculate long-term trends in energy consumption. The residents should be able to configure the minimal width of the aggregation window. We propose an aggregation window of one day to be a good balance between privacy and usefulness of the aggregated values.

[2] https://discovergy.com

[3] https://plotwatt.com

2.2 Access Control Rules

In this section, we define the $read$ operations of the database which enforces our access control mechanism. There are two $read$ operations, one to access fine-grained data and one for aggregated data. The $read$ operations should meet all requirements above. Timestamps are assumed to be POSIX times with resolution in seconds.

- Let U be the set of users / residents of the monitored environment.
- Let D be the set domains within the environment.
- Let T be the set of all sensor value time series.
- Each time series is a function mapping from a timestamp (as integer value) to a sensor value whereby there might be missing data ($\perp$):
$$t_i \in T : \mathbb{Z} \to \mathbb{R} \cup \{\perp\}.$$
- The function $d : T \to D$ associates each time series to a specific domain.
- Special time series record the presence of a particular resident in a domain:
$$t_{(u,d)} : \mathbb{Z} \to \{\top, \perp\} \text{ with } u \in U \text{ and } d \in D.$$

Now the basic $read$ operation on a time series t_i at time τ can be defined based upon whether the presence of the user u currently requesting the data was detected in the domain t_i is associated with:

$$read(t_i, \tau) = \begin{cases} t_i(\tau), & t_{(u,\ d(t_i))}(\tau) = \top \\ \perp, & otherwise \end{cases}.$$

With this definition, in the time periods where no presence in a specific domain was recorded, all the time series of this domain are displayed as if the data is missing. Thus, requirement 1 is met. Fig. 1 shows the effect of this definition under the assumption that the user was detected absent during the time periods marked purple.

Time series aggregation needs to be defined in order to express requirement 3:

- Let O be the set of aggregation operations (e.g. sum, max, or $mean$).
- $A \subseteq T \times \mathbb{Z} \times O$ is the set of aggregated time series which is used by the visualization system. Each aggregated time series applies a specific operator $o_j \in O$ to a time series in $t_i \in T$ with a time window size $r \in \mathbb{Z}$.
- Let $r_{disclose}$ be the minimal window where all time series data can be accessed by all valid users U (e.g. all residents). In Requirement 3b, we propose $r_{disclose} = 86400$ seconds which is one day.

The read operation on aggregated time series (t_i, r, o_j) at time τ is then defined as:

$$read\big((t_i, r, o_j), \tau\big) = \begin{cases} \perp & , \quad \tau \bmod r \neq 0 \\ o_j(\{t_i(l) : l \in [\tau, \tau + r - 1]\}) & , \quad \begin{array}{c} r \bmod r_{disclose} = 0 \ \vee \\ \exists k \in [\tau, \tau + r - 1] : t_{(u,d(t_i))}(k) = \top \end{array} \\ \perp & , \quad otherwise \end{cases}$$

The first case forces that the requested time τ is aligned to the window r to avoid leaking information about the high resolution time series. In the second case, the aggregated data is provided either if the requested window is a multiple of the minimal window ($r \bmod r_{disclose} = 0$) or the user was present in the domain $d(t_i)$ at least one second. The constraint that r needs to be a multiple of $r_{disclose}$ is also due to avoid possible information leakage.

With these two definitions of the $read$ operation, both requirement 1 and 3 can be met. The bottom row of Fig. 1 illustrates that the average daily consumption of each device can be accessed independently of presence. Requirement 2 is a technical issue which is discussed in the next section.

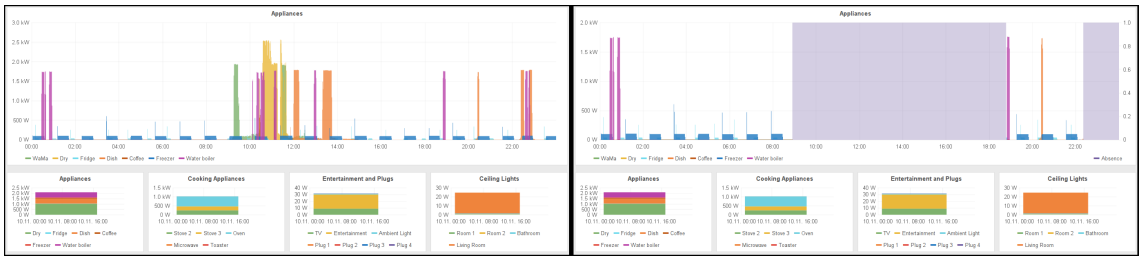

Figure 1 : Visualization systems (left frame) display sensitive information about actions inside the household. In the right frame, data recorded during the absence (marked purple) of the resident is hidden from him.

2.3 Tracking and displaying the presence

The proposed system can only be implemented when a person's presence can be detected in each domain. A practical technology is based on Bluetooth LE beacons which can be attached to key chains. These beacons constantly transmit an identification for the owner. Passive Bluetooth receivers located in each domain can use this identification to assume presence of a particular person which is then stored in the database.

There is no guarantee that the residents will always carry around their identification token. In fact, someone could maliciously hide his token in a domain to gain access to the recorded data later. This threat can be dealt with by displaying all detected presences on a display. Each person can then check whether they are really alone.

3 CONCLUSION AND FUTURE WORK

We presented a concept of a monitoring and visualization system for energy and sensor data, which implements mechanisms to avoid surveillance of fellow residents or co-workers. However a secure implementation requires further investigations. For example, the proposed Bluetooth beacons used for identification can easily be spoofed. We need a non-replayable identification mechanism using cryptographic primitives. Another question is how to handle situations where a resident is accountable for consuming energy, but is not present at the time of the actual consumption. This occurs when appliances are programmed to run later or when appliances are controlled remotely.

3.1 Acknowledgement

This interdisciplinary work was funded in the KASTEL project by the German Federal Ministry of Education and Research (BMBF 01BY1172).

REFERENCES

[1] Allerding, F., Mauser, I., & Schmeck, H. (2014). *Customizable Energy Management in Smart Buildings Using Evolutionary Algorithms*. In A. I. Esparcia-Alcázar & A. M. Mora (Eds.), Applications of Evolutionary Computation (pp. 153–164). Springer Berlin Heidelberg.

[2] Molina-Markham, A., Shenoy, P., Fu, K., Cecchet, E., & Irwin, D. (2010). *Private memoirs of a smart meter*. In Proceedings of the 2nd ACM workshop on embedded sensing systems for energy-efficiency in building (pp. 61–66). ACM.

[3] Lisovich, M. A., Mulligan, D. K., & Wicker, S. B. (2010). *Inferring Personal Information from Demand-Response Systems*. IEEE Security Privacy, 8(1), 11–20.

[4] Decision of the German Federal Constitutional Court in BVerfGE Vol. 65, p.1.