Karlsruher Institut für Technologie

Institut für Programmstrukturen und Datenorganisation (IPD)
Lehrstuhl für Systeme der Informationsverwaltung, Prof. Böhm

Bachelor/Master Thesis

Bachelor/Master Thesis

Measuring the Privacy Loss with Smart Meters

Smart Meters will play a significant role in future electricity systems. At the same time, it is widely recognized that fine-grained electricity consumption data (a *load profile*) provided by these devices contains private information or *secrets* about households equipped with them. Examples of such secrets often used in literature are:

List 1

- Vacancy patterns
- Habits
- Activities
- Appliance usage
- Sleep-wake cycles

and many others. Some studies conclude that even the TV channel being watched can be discovered from such data.

One of the proposed techniques to mitigate the privacy loss is to modify a load profile with the use of a rechargeable battery. There are several tens of such battery charging algorithms (*BCA*) proposed for this purpose so far.

The quality in terms of gained privacy by BCA is computed by comparing a load profile with and without using the algorithm. This function used to compare the profiles is a *privacy measure*. There are several privacy measures in the literature. Examples include:

List 2

- cross-correlation
- mutual information
- R-squared
- entropy

and many more. **However, it is unclear to what extent these measures are able to capture whether a specific secret is hidden in the data or not**. **That is, if a privacy measure from list 2 matches the secrets in list 1**. For example, can one infer from the value of 0.5 for cross-correlation that the sleep-wake cycles are not detectable using a certain BCA?

The task of this work is to find answers to the general connection between secrets from list 1 and privacy measures from list 2. To do so, the following steps are important:

1. For a selected set of secrets develop formalizations of when the secret is hidden
2. For these secrets develop a test which checks if a privacy measure is able to distinguish cases reliably, when the secret is hidden, and when it is not.
3. Run these tests on the privacy measures and draw the conclusions.

Throughout this work, you will develop strong R programming skills, and get familiar with concepts both from data mining and information theory. This project also implies a creative approach as the goal is stated precisely, but there is a lot of freedom in choosing the exact methods.

Contact

Vadim Arzamasov vadim.arzamasov@kit.edu Room: 340

Am Fasanengarten 5 76131 Karlsruhe Building: 50.34