

TWO IS NOT ENOUGH

PRIVACY ASSESSMENT OF AGGREGATION SCHEMES IN SMART METERING

Niklas Büscher, **Spyros Boukoros**, Stefan Bauregger, Stefan Katzenbeisser
Technische Universität Darmstadt

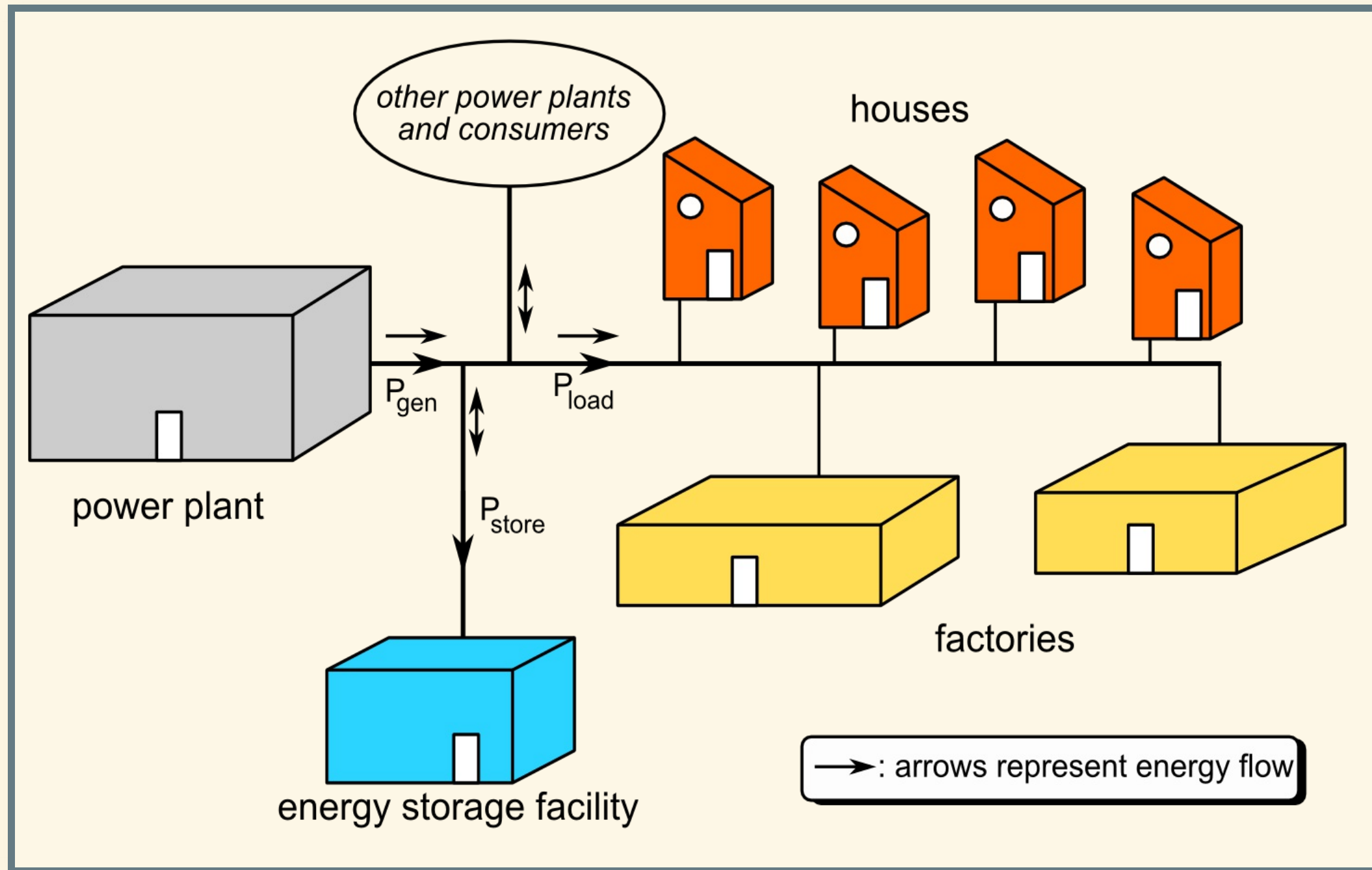


Privacy and Trust
for Mobile Users

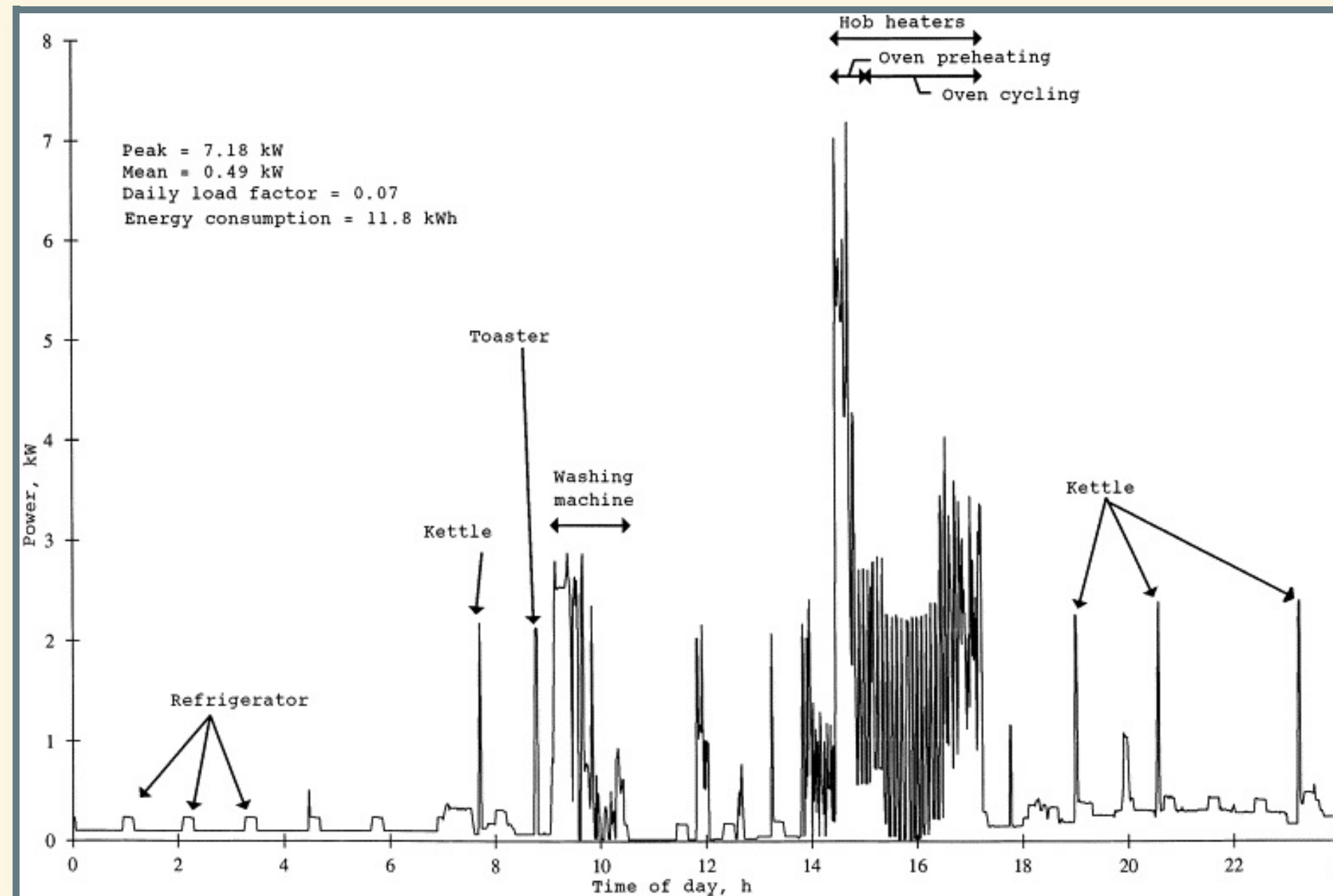


TECHNISCHE
UNIVERSITÄT
DARMSTADT

POWER GRID



Example load profile



G. Wood and M. Newborough. Dynamic energy-consumption indicators for domestic appliances: environment, behaviour and design. *Energy and buildings*, 35(8):821–841, 2003.

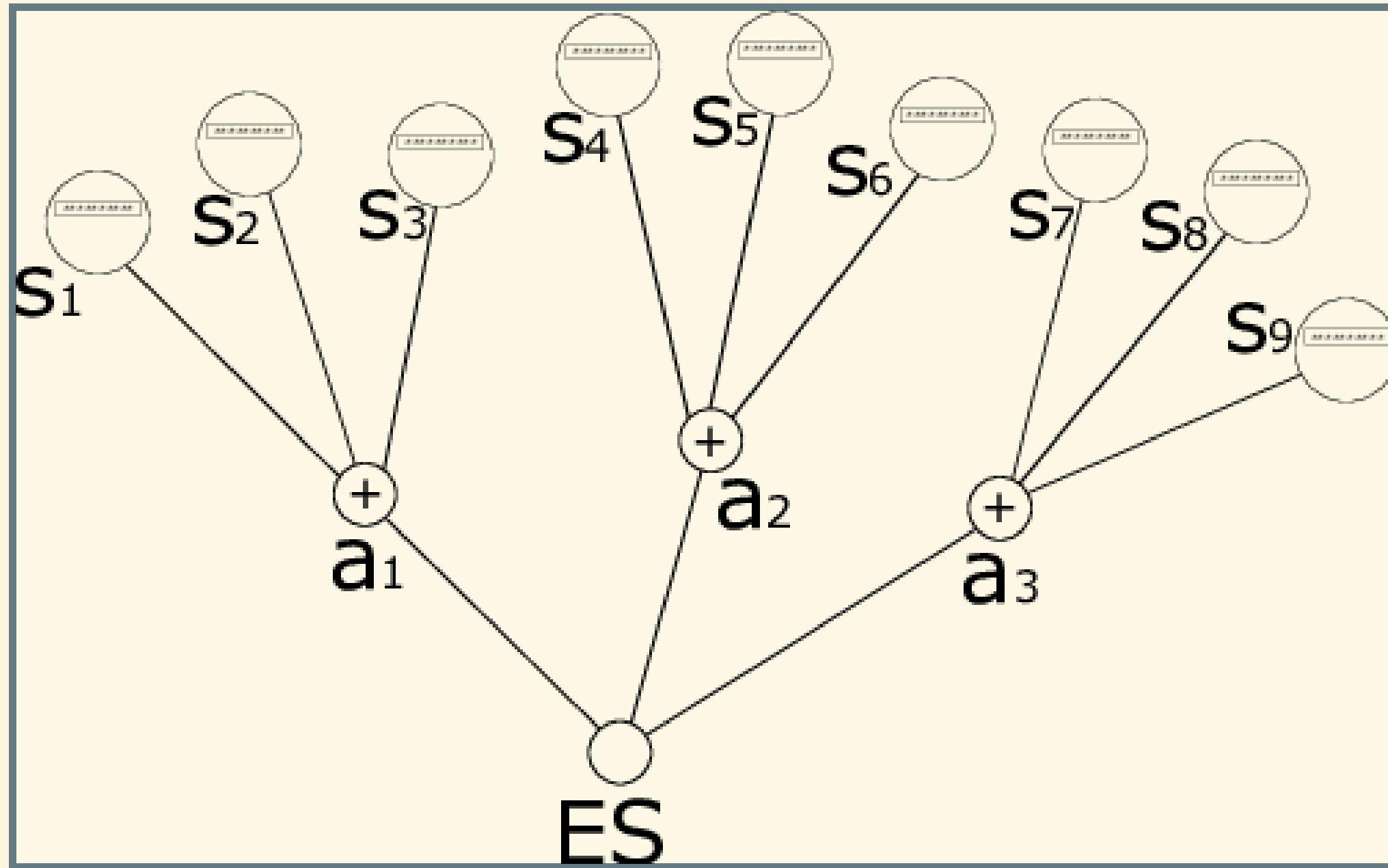
AGGREGATION SCHEMES

is a prominent solution

Instead of reporting individual households

Report many together

AGGREGATION EXAMPLE



What is the minimum # of households necessary ?

*This report suggests that selecting an **aggregation level of 2** offers network companies greater visibility ... while still providing customers with a comparatively similar level of visibility risk to an aggregation level of 4*
— Energy Networks Association (2015 report)

Let's **test** this!

We base our metric on the notion of **indistinguishability**

J. M. Bohli, C. Sorge, and O. Ugus, “A privacy model for smart metering,” in 2010 IEEE International Conference on Communications Workshops, pp. 1–5, May 2010.

PRIVACY GAME

Inputs: E_{gen}, m, l

Adversary

$\hat{e}_0, \hat{e}_1 \leftarrow E_{gen}$

Challenger

$\xrightarrow{\hat{e}_0, \hat{e}_1} r \leftarrow_{\$} \{0, 1\}$

$\hat{e}_2, \hat{e}_3, \dots, \hat{e}_m \leftarrow_{\$} E_{gen}$

compute aggregate $\hat{e}_a$:

$\xleftarrow{\hat{e}_a} \hat{e}_a \leftarrow \hat{e}_r + \hat{e}_2 + \dots + \hat{e}_m$

$g \leftarrow f_{dec}(\hat{e}_a, \hat{e}_0, \hat{e}_1) \xrightarrow{g} \text{output } (g = r)$

METHODOLOGY

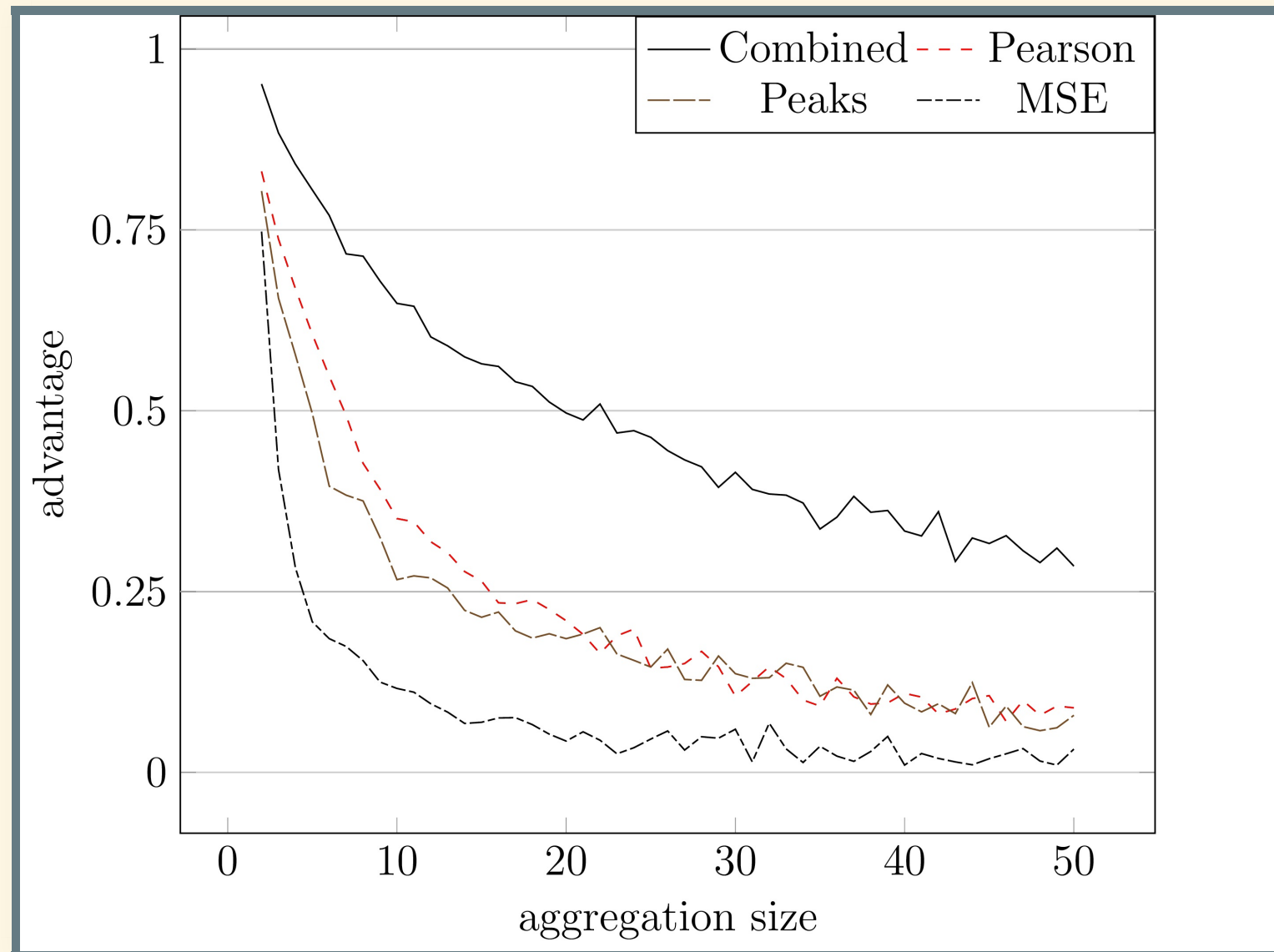
- Measure advantage over random guessing
- Rely on simple heuristics
 - Peak comparison
 - Mean square error
 - Pearson correlation
 - **Combined method** based on
Peak comparison and the Pearson correlation
- Real world data
 - Largest available datasets

WE EXAMINE THE FOLLOWING CASES:

- Can we distinguish daily load profiles in the aggregate?
- Can we distinguish devices in the aggregate?
- Parameters affecting our game?

Are small aggregations privacy preserving? (1/3)

One dataset - Daily load profiles

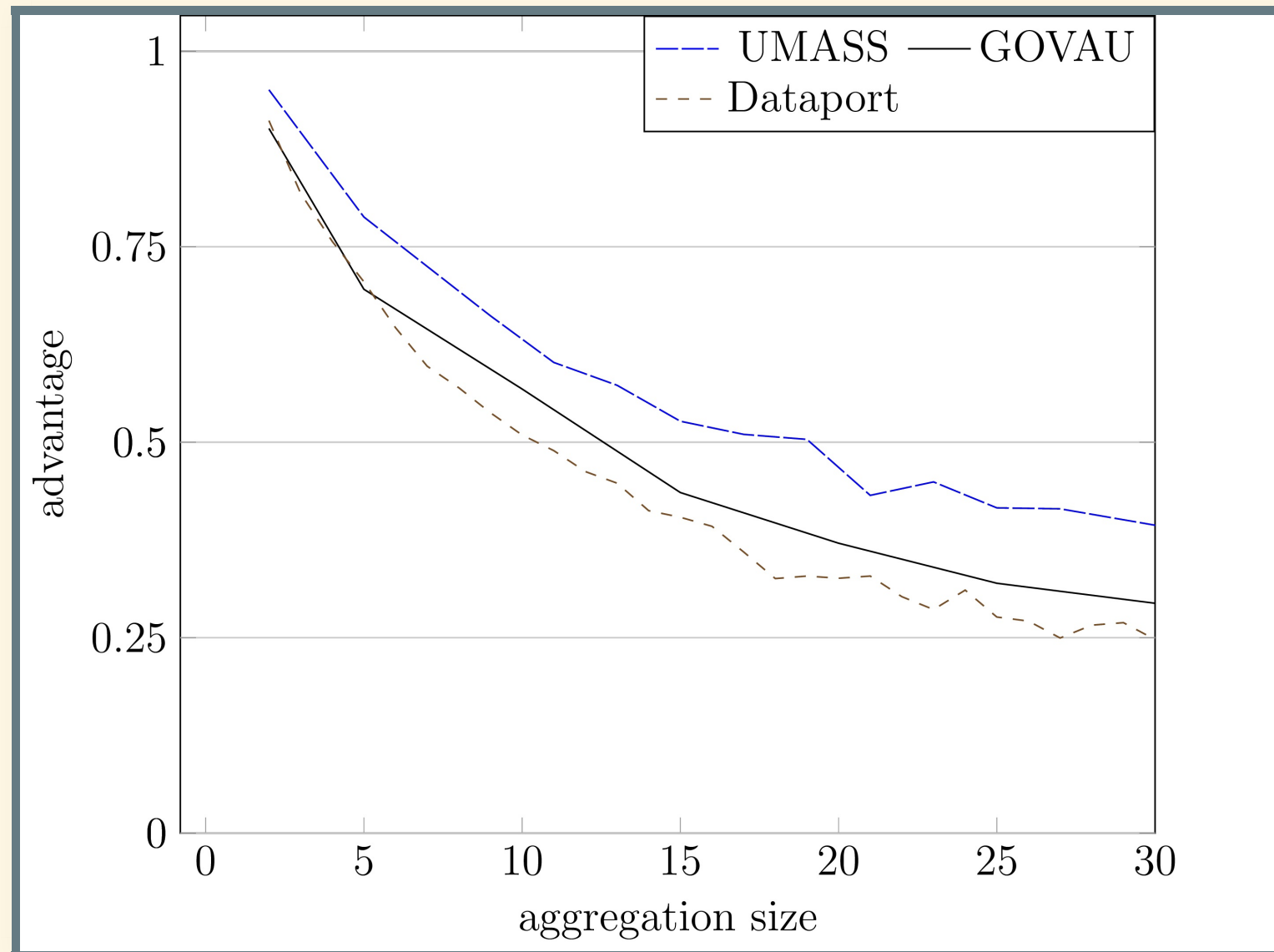


- No privacy in small aggregations
- Significant advantage in larger aggregations

Dataport dataset with 15 minutes resolution.

Are small aggregations privacy preserving? (2/3)

Multiple datasets - Daily load profiles

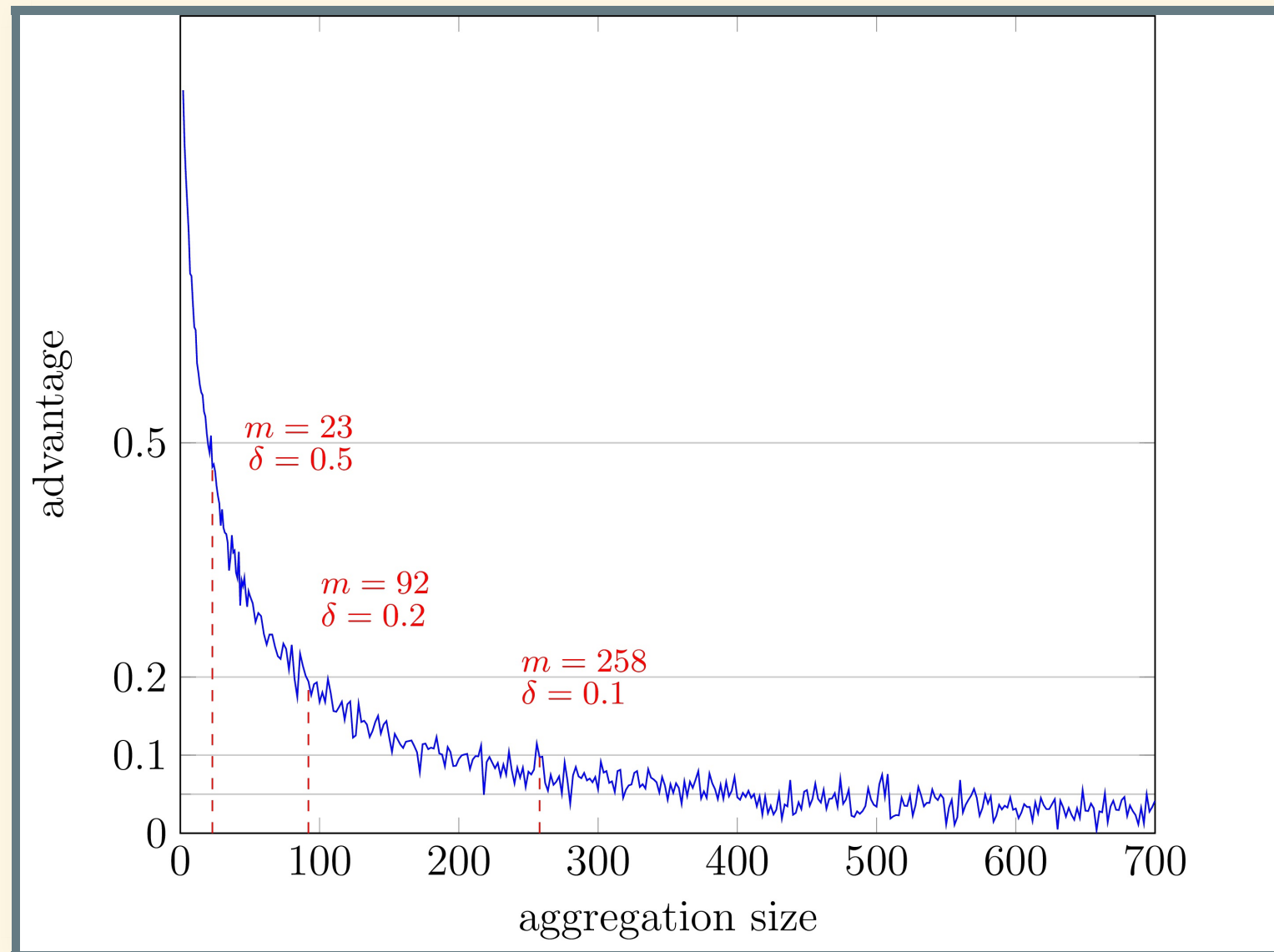


- Different datasets:
- Geographical origin
 - # of households
 - # of load profiles per household
 - Types of energy

Multiple datasets-30 minutes resolution.
Combined method.

Are small aggregations privacy preserving? (3/3)

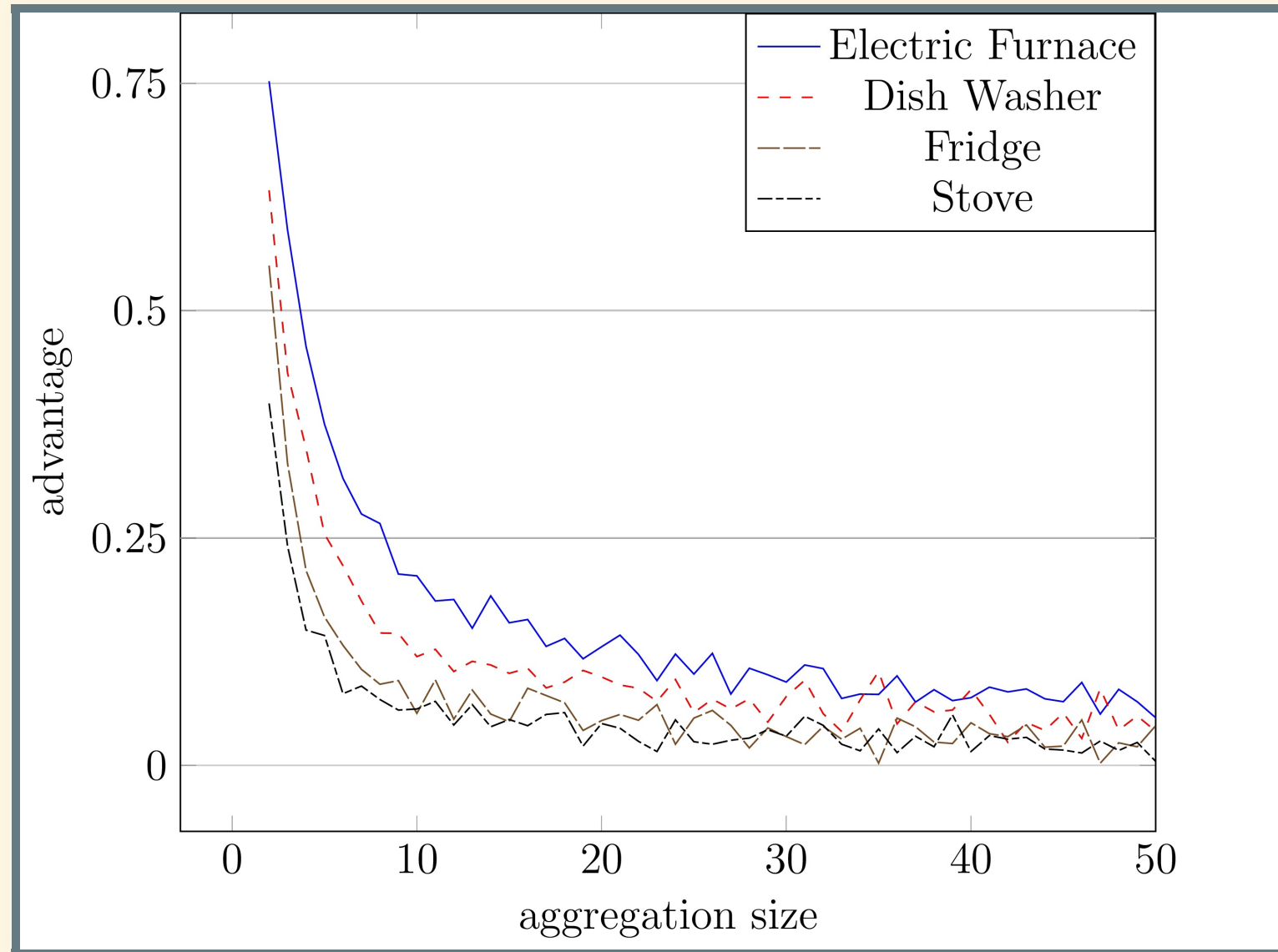
Privacy limit - Daily load profiles



- Heavily depends on the dataset
- Larger aggregations introduce a lot of noise
- More research regarding utility vs privacy

Dataport dataset with 15 minutes resolution.
Combined method.

Are single appliances detectable in the aggregate?

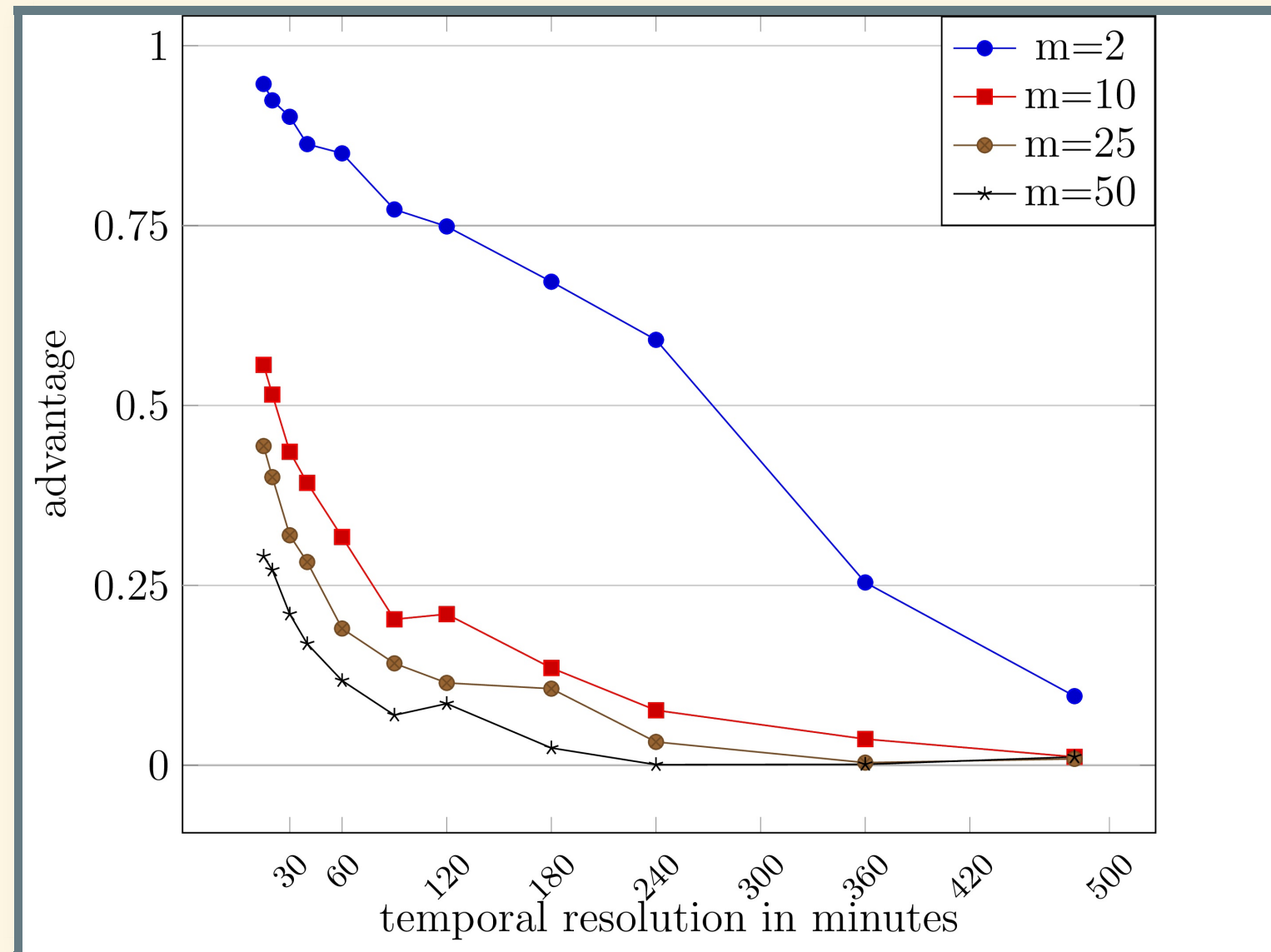


- Modified privacy game with/without device
- Energy hungry appliances easier to detect
- More devices in the paper
- Correlation of device characteristics and detectability

Dataport dataset with 15 minutes resolution.
Combined method.

Parameters affecting the privacy game

Temporal resolution



- Less freq. reports more private
- No privacy in small agg.sizes
- More parameters in the paper

Dataport dataset using the combined method.

IN A NUTSHELL..

- Small aggregations cannot guarantee privacy
 - Individual profiles
 - Single devices
- An upper limit seems to exist but...
 - dataset dependent
 - privacy vs (meaningful) utility
- Temporal resolution is an important factor

Two (or just a few) is **definitely** not enough!

Thank you!

