Learning the Energy Consumption Behavior of Users by Clustering

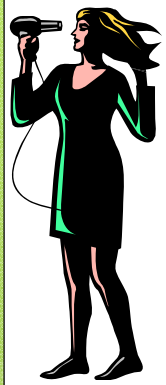
Michela Milano, Marco Patella, Cristina Lelli

DISI – Università di Bologna

Our role in the project

- Analyze the huge amount of data derived from devices installed in houses of trial users

- User Profiling with the goal of:
  - Advising
  - Appliances scheduling
  - Reverse engineering of energy consumptions



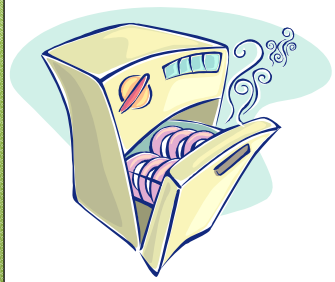
User profiles

- Starting from data readings coming from “smart appliances” and “smart plugs”
  - Derive a user profile
  - Highlighting (possibly existing) habits
  - According to the existence of “similar” energy consumptions
    - Daily
    - Weekly
    - Monthly
- Users can then be grouped according to the “similarity” of relative profiles

The user profile

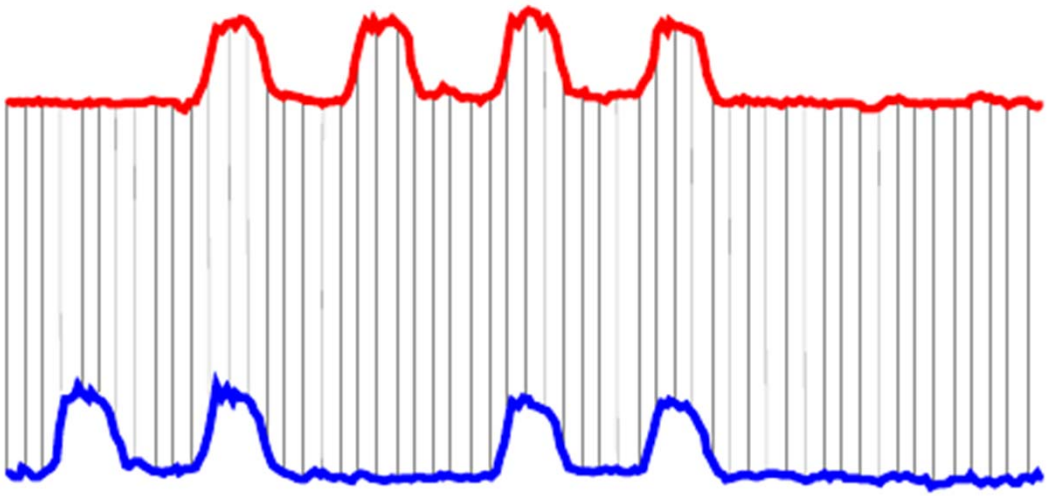
- The profile of a user can be useful for:
  - Suggesting the user more "appropriate" habits (e.g., for saving money)

  - Scheduling the activation of some appliances, according to forecasted consumptions and/or to the presence of micro-generation



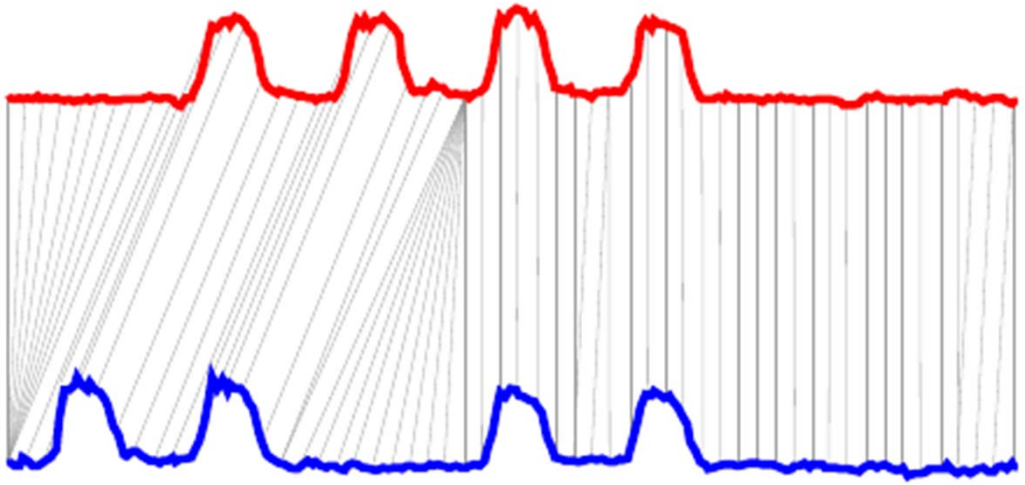
  - In case smart appliances or plugs are not available, modeling the usage of appliances, by analyzing current energy consumption

How user profiles are obtained



- Comparison between daily/weekly/monthly energy consumptions, yielding a similarity value
- This can be done by comparing energy consumption at each time interval
- This is very sensitive to the presence of "time shifts" in the consumptions

How user profiles are obtained



- By using Dynamic Time Warping, we can grade as similar two series, even if they are non-perfectly aligned
- We can control the amount of “warping”, e.g. to avoid matching “distant” consumptions

Energy@home

How user profiles are obtained

- A habit is a sufficiently large group of sufficiently similar consumptions (DBScan)
- We can control the amount of both "sufficiently"
- Advantages:
  - General application (we can change the notion of similarity maintaining the clustering algorithm)
  - Simplicity of parameters (e.g. number of clusters is an output, not an input)

How user profiles are obtained

- Disadvantages:
  - A cluster could include series that are (very) different if they are both similar to a third series (and so on)
- We form clusters only considering series that are all sufficiently similar to each other (a completely connected subgraph in the connectivity graph)
- This does not define a partition
  - Acceptable, since an user could be part of several “groups” of users

What's next?

- We (deliberately) ignore demographic aspects for clustering purposes
  - Such information can be used a-posteriori, e.g. to derive characteristics shared by users in a single cluster
  - This could help in classifying a new user, for which we do not have any consumption data, without the need for training the system