GREEND Parser and Metadata for NILMTK

Johannes Hölzl
The Dark Side of Force
johannes.hoelzl@gmail.com

Abstract—**GREEND[MEE⁺14] is a dataset which has been developed and collected in Carinthia and parts of Italy. It contains eight buildings and a couple of months of power measurement.**

NILM-Metadata[KK14] is project which was designed for being used with NILM-TK[BKP⁺14], the free and open source Non-Intrusive Load Monitoring Toolkit.

I. INTRODUCTION

The GREEND parser has been developed in the thought of converting this particular dataset into a general, easy-to-use and most of all, a comparable structure. As many other parsers for NILM-TK, it converts the dataset into a standardized format which is also known as NILMTK-DF. Using a hdf5 file instead of a punch of `csv` or similar files has advantages such as the increase of performance when importing the dataset to the program, overall size on disk, which is much smaller and the circumstance that is just a single file containing all data and metadata makes it much more comfortable to use than the previously described scenario.

II. USING THE PARSER

The GREEND parser is stored in a file which is called `convert_greend.py` as the function `convert_greend`. It takes to strings as parameters which both describe paths. The first parameter gives in the path to the directory in which the GREEND dataset is stored. The second one is the location in which the outputted hdf5 file should be saved to.

When using the parser, keep in mind that GREEND is using up quite some disk space because of its sample rate of one second and its long recording period.

A. Example on using the GREEND parser

This Example shows the Usage of the GREEND parser which converts the data stored at `greend/path` to a hdf5 file at `path/to/greend.h5`.

```
convert_greend('greend/path',
    'path/to/greend.h5')
```

Fig. 1. Example of Converting GREEND

B. Using the Dataset

The hdf5 file could now be used with NILM-TK as expected. With NILM-TK v0.2 there is a class called `HDFDataStore` which could be, and is usually used to import hdf5 files. To finally load the dataset in form in which it could be used, it is indispensable to create a `DataSet` object and call its function `load()` with the `HDFDataStore` as a parameter.

C. Technical Background

The GREEND converter is not much different than the REDD one. They differ in the fact, that every meter contained in REDD has its own file with data in it, while in GREEND, every day is stored in a separate file with multiple columns for the appliances.
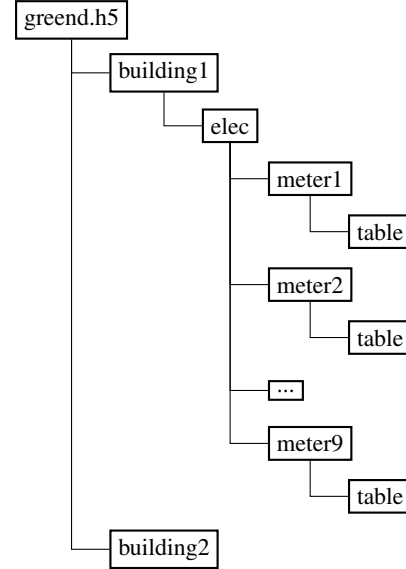

Fig. 2. Structure of the Dataset

REFERENCES

[BKP⁺14] Nipun Batra, Jack Kelly, Oliver Parson, Haimonti Dutta, William Knottenbelt, Alex Rogers, Amarjeet Singh, and Mani Srivastava. Nilmtk: An open source toolkit for non-intrusive load monitoring. In *Proceedings of the 5th International Conference on Future Energy Systems*, e-Energy '14, pages 265–276. ACM, 2014.

[KK14] Jack Kelly and William Knottenbelt. Metadata for Energy Disaggregation. In *The 2nd IEEE International Workshop on Consumer Devices and Systems (CDS 2014)*, Västerås, Sweden, July 2014.

[MEE⁺14] Andrea Monacchi, Dominik Egarter, Wilfried Elmenreich, Salvatore DAlessandro, and Andrea M. Tonello. GREEND: an energy consumption dataset of households in Italy and Austria. In *Proceedings of the IEEE International Conference on Smart Grid Communications (SmartGridComm'14)*, November 2014.