Two approaches for synthesizing scalable residential energy consumption data

Many fields require scalable and detailed energy consumption data for different study purposes. However, due to privacy issues, it is often difficult to obtain sufficiently large datasets. This paper proposes two different methods for synthesizing fine-grained energy consumption data for residential households, namely a regression-based method and a probability-based method. They each use a supervised machine learning method, which trains models with a relatively small real-world dataset and then generates large-scale time series based on the models. This paper describes the two methods in details, including data generation process, optimization techniques, and parallel data generation. This paper evaluates the performance of the two methods, which compare the resulting consumption profiles with real-world data, including patterns, statistics, and parallel data generation in the cluster. The results demonstrate the effectiveness of the proposed methods and their efficiency in generating large-scale datasets.

General information
Publication status: Published
Organisations: Sustainability, Energy Systems Analysis, Department of Technology, Management and Economics, Department of Civil Engineering, Energy and Services, University College of Northern Denmark, University of Technology Sydney
Corresponding author: Liu, X.
Contributors: Liu, X., Iftikhar, N., Huo, H., Li, R., Nielsen, P. S.
Pages: 586-600
Publication date: 2019
Peer-reviewed: Yes

Publication information
Journal: Future Generation Computer Systems
Volume: 95
ISSN (Print): 0167-739X
Ratings:
BFI (2019): BFI-level 1
Web of Science (2019): Indexed yes
Original language: English
Keywords: Energy consumption, Time series, Synthesize, Simulation, Data generation
DOIs:
10.1016/j.future.2019.01.045
Source: FindIt
Source ID: 2443610223
Research output: Contribution to journal › Journal article – Annual report year: 2019 › Research › peer-review