Energy 159 (2018) 302–309

Contents lists available at ScienceDirect

Energy

journal homepage: www.elsevier.com/locate/energy

An analysis of utility meter data aggregation and tenant privacy to support energy use disclosure in commercial buildings

Olga V. Livingston, Trenton C. Pulsipher, David M. Anderson, Alex Vlachokostas, Na Wang*

Pacific Northwest National Laboratory, 902 Battelle Blvd, Richland, WA, 99354, USA

A R T I C L E I N F O

Article history:
Received 16 January 2018
Received in revised form
6 June 2018
Accepted 20 June 2018
Available online 21 June 2018

Keywords:
Energy disclosure
Commercial buildings
Tenant privacy
Meter aggregation
Utility policies
Benchmarking

A B S T R A C T

A growing number of cities are adopting energy use benchmarking ordinances, which require building owners report their buildings' total energy usage annually. It also requires that utilities supply aggregated building-level monthly energy consumption data. The data aggregation poses privacy concerns, as it is possible to estimate the individual tenant load consumption curve by dividing aggregated energy data by the number of meters. A solution is to quantify and assess the impact of adjusting the utility meter aggregation threshold on tenant privacy and on buildings that are eligible for energy usage reporting. As the threshold increases, fewer buildings are eligible for energy use data disclosure and therefore lessening data value. This study aims to investigate the similarity between individual utility meters and whole-building totals at various aggregation levels. Based on statistical analysis of 715,000 anonymized, non-residential meter accounts from six utilities across the U.S., we developed the "Meter Aggregation Selection Threshold" as a metric to assess tenant privacy risk. The metric estimates the portion of individual customer energy use patterns that are similar to the aggregated building consumption profile. It allows policy makers to make an informed decision on whether required disclosure regulations compromise business sensitive information and tenant privacy.

1. Introduction

Benchmarking is a process in which building performance is evaluated by comparison with a baseline or peers. The aim is to improve energy efficiency, minimize energy consumption, and enable transparency in energy use of buildings. Regulatory bodies, through building regulation, energy auditing [1], and energy rating [2] promote and support energy efficiency investment projects [3]. As a result, more building owners are adopting benchmarking as a driver of energy and cost savings. Benchmarking enables owners to assess buildings' energy consumption, and identify cost-effective energy upgrades. To perform benchmarking, the building owner must provide information regarding the building's physical characteristics and very importantly, monthly whole-building energy usage information. The main barrier to benchmarking lies within the process of acquiring monthly whole-building energy usage information. This is a challenging task for commercial buildings, where individual tenants own their utility meters and accounts. The building owner relies on the utility to obtain aggregated building energy usage information. This information is then used as an input to benchmarking tools for an energy consumption assessment, and to identify opportunities for energy efficiency measures. States across the US [4] have legislated commercial building benchmarking and disclosure regulations under which utilities are mandated to facilitate the provision of energy use data to help building owners comply with benchmarking requirements.

The benchmarking and energy disclosure policies require the building owners to measure and disclose their energy use. These aim to reduce the building sector's energy use by first assessing inefficient building energy assets and consequently implementing energy-efficiency projects. Meng et al. [5] estimated that benchmarking and disclosure policies in NYC reduced energy use by 6% after the first three years of the policy implementation and 14% after four years. Also, Lee et al. [6,7] implemented benchmarking as a tool for energy management of cooling processes in buildings. Nevertheless, benchmarking is not limited to the building sector but has been applied to various industrial processes. For example, Iribarren et al. [8] used benchmarking analysis for the wind power sector. Cai et al. [9] created a tool based on energy benchmarking

* Corresponding author.
E-mail addresses: olga.livingston@hq.dhs.gov (O.V. Livingston), trenton.pulsipher@henryschein.com (T.C. Pulsipher), DMA@pnnl.gov (D.M. Anderson), alex.vlachokostas@pnnl.gov (A. Vlachokostas), nora.wang@pnnl.gov (N. Wang).

https://doi.org/10.1016/j.energy.2018.06.133
0360-5442/

for assessing energy demand and efficiency of machining systems. Mirzakhani et al. [10] used energy benchmarking with the goal of reducing energy consumption in the cement industry sector, and lastly Gilbert et al. [11] benchmarked the natural gas and coal-fired electricity generation in the US. All the aforementioned studies relied on energy data to perform benchmarking.

Utilities face a paradoxical situation due to contradictory rules. On the one hand, utilities are required to provide building owners with energy usage data to comply with benchmarking requirements. On the other hand, when releasing energy usage data, utilities risk data privacy violations. The State & Local Energy Efficiency (SEE) Action Customer Information and Behavior Working Group has defined three categories of Customer Energy Use Data (CEUD) [12]. The first category consists of Personally Identifiable Information (PII), such as names, addresses, and social security numbers. The second category is composed of customer-specific energy usage data such as total- and time-differentiated energy uses. Finally, the third category includes aggregated data that the utility assembles from multiple meters, to provide information about energy consumption across a specified area. The present paper focuses on the data from the third category, specifically used for benchmarking purposes through ENERGY STAR Portfolio Manager (ESPM) benchmarking tool [13]. Data aggregation from a cybersecurity perspective is not discussed in this paper, but cannot be found in our full technical report [14]. Current efforts to justify a standard meter aggregation threshold are limited and most utilities follow arbitrary or empirical rules. Our study, serves as a reference for current practitioners with a statistical-based justification on the number of meters that can be safely aggregated without revealing individual tenant energy usage consumption profiles. The strength of the present study is that its methodology can be applied to any meter aggregation effort and act as a statistically justifiable metric for comparing other aggregation thresholds. As a result, utilities can compare their aggregation threshold against a common metric.

Currently, utilities provide Customer Energy Use Data (CEUD) [15] without following a standard framework or ruleset. The utilities aggregate CEUD to provide whole-building energy usage data to building owners. The intent is to protect customer's privacy because it is belived that it is difficult to estimate the individual tenant's energy usage when meters are aggregated. There are three simple concepts behind CEUD aggregation and release that we define as follows:

- Utility meter aggregation: Utility meter profiles in a building are aggregated (i.e., summed) to form a total monthly energy usage profile across all utility meters associated with a building. In this paper, we group the buildings based on the number of utility meters that are present at each building (e.g., an "N-meter building," assumes N utility meters present in the building.)
- Aggregation threshold: We define aggregation threshold as the minimum number of utility meters that are aggregated in a building without privacy concerns for the tenants. For example, when aggregating meters in a building, the average building meter profile (ABMP) curve may be similar to the meter profile of individual tenants. We define ABMP as the sum of the meter consumption of a building over the total meters in the building.
- Number of buildings covered: The choice of a specific aggregation threshold will have a direct impact on the type of buildings that will be covered by the threshold. For example, typically there are more buildings that have four meters rather than five meters or more. As a result, raising the aggregation threshold, from 4- to 5-m buildings, results in fewer buildings since a large number of buildings are excluded from the pool.

In Fig. 1 we illustrate the aforementioned concepts. Fig. 1(a) shows that in a utility's buildings sample the number of buildings equipped with a large number of utility meters is considerably less than the number of buildings with a small number of utility meters. In addition, Fig. 1(b) shows qualitatively, the exponential decrease of the utility meter percentage that resembles the Average Building Meter Profile (ABMP) as the aggregation threshold increases. Therefore, in buildings with fewer utility meters, individual tenant meter readings would likely coincide with the ABMP. The ABMP is calculated by dividing the aggregated meter data of the building by the number of utility meters. This re-identification (or back-calculation) is not desirable because it entails privacy concerns. As a result, we are faced with a Pareto efficiency type of problem, meaning that when optimizing for one variable (i.e., report as many buildings in benchmarking) then another variable is negatively affected (i.e., cause privacy concerns). We notice that effect when trying to preserve tenant user data privacy and at the same time release benchmarking data from as many users as possible. Results at Section 4 show that, by changing the threshold of meter aggregation, the meter profiles that are similar to the average meter profile change as well.

The aim of the paper is to analyze a utility meter profile aggregation methodology with emphasis on the following questions:

A. What is the probability that energy usage of an individual utility meter can be back-calculated if the total number of meters within the building, and total building energy consumption, are known?
B. What percentages of individual utility meters have consumption values that resemble ABMP?
C. As the aggregation threshold increases, how does this impact the number of buildings eligible to receive whole-building aggregated data?

The paper is organized as follows: In Section 2, we review the existing literature regarding privacy and identification of individuals through energy use data. In Section 3, we present our analytical methodology and discuss how our study contributes to the field. In Section 4, we present the results of our methodology. Finally, Section 5 contains our study conclusions and recommendations for utilities and practitioners on how to safely aggregate and release energy use data.

2. Existing research and current practice

We start the literature review by discussing current utility practices. Currently, utilities follow rules of thumb, which are not justified by any statistical model. More specifically, some utilities use the 4/80 aggregation rule, requiring them to aggregate energy

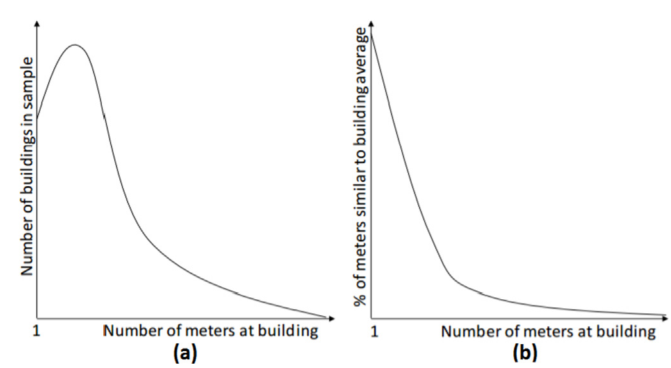

Fig. 1. Qualitative relationship showing (a) the number of utility meters in the buildings sample and (b) the impact of aggregating more utility meters on the ABMP.