

Diss. ETH No. 23857

Screening Meter Data: Characterization of Temporal Energy Data from Large Groups of Non-Residential Buildings

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES of ETH ZÜRICH
(Dr. sc. ETH Zürich)

presented by

CLAYTON C. MILLER

Masters of Science (MSc.) Building, National University of Singapore

Masters of Architectural Engineering (MAE), University of Nebraska

born on 22 January 1984

citizen of the United States of America

accepted on the recommendation of
Prof. Dr. Arno Schlueter, examiner
Prof. Dr. Stefan Mueller Arisona, co-examiner

2016

Abstract

This study focuses on the screening of characteristic data from the ever-expanding sources of raw, temporal sensor data from commercial buildings. A two-step framework is presented that extracts statistical, model-based, and pattern-based behavior from two real-world data sets. The first collection is from 507 commercial buildings extracted from various case studies and online data sources from around the world. The second collection is advanced metering infrastructure (AMI) data from 1,600 buildings. The goal of the framework is to reduce the expert intervention needed to utilize measured raw data in order to extract information such as building use type, performance class, and operational behavior. The first step is feature extraction and it utilizes a library of temporal data mining techniques to filter various phenomenon from the raw data. This step transforms quantitative raw data into qualitative categories that are presented in heat map visualizations for interpretation. In the second step, or the investigation, a supervised learning technique is tested in the ability to assign impact scores to the most important features from the first step. The efficacy of estimating variable causality of the characterized performance is tested to determine scalability amongst a heterogeneous sample of buildings. In the first set of case studies, characterization as compared to a baseline was three times more accurate in characterizing primary building use type, almost twice for performance class, and over four times for building operations type. For the AMI data, characterizing the standard industry class was improved by 27% and predicting the success of energy savings measures was improved by 18%. Qualitative insight from several campus case study interviews are discussed as well. The usefulness of the approaches was discussed in the context of campus building operations.

Kurzfassung

Diese Studie behandelt das Sichten, Sortieren und Bearbeiten charakteristischer Zeitreihen aus stark wachsenden Quellen für rohe Sensordaten in kommerziellen Gebäuden. Ein zweistufiges Vorgehen wird präsentiert, das statistische, modellbasierte und Musterbasierte Verhaltensweisen von zwei Datensätzen extrahiert. Der erste Datensatz beinhaltet Daten von 507 kommerziellen Gebäuden, zusammengetragen aus verschiedenen Fallbeispielen und online Datenquellen aus der ganzen Welt. Der zweite Datensatz beinhaltet Daten von Advanced Metering Infrastructure (AMI) von 1,600 Gebäuden. Das Ziel der vorgestellten Methode ist das Reduzieren benötigter Experteneingriffe, um gemessene Rohdaten benutzen zu können zum Erhalten von Information wie Gebäudenutzungstyp, Performance Klasse und Betriebsverhalten. Im ersten Schritt, dem Extrahieren von Charakteristiken, werden durch das Benutzen einer Bibliothek von Data Mining Techniken verschiedene Phänomene aus den Rohdaten herausgefiltert. Dieser Schritt transformiert quantitative Rohdaten zu qualitativen Kategorien, die durch Heat Map Visualisierungen präsentiert und interpretiert werden. Im zweiten Schritt, der Datenuntersuchung, wird eine Supervised Learning Technique auf die Möglichkeit hin getestet, den wichtigsten Charakteristiken aus dem ersten Schritt eine Auswertung der Auswirkungen zuzuordnen. Um das Hochskalieren für heterogene Gebäudeparks zu untersuchen wird die Wirksamkeit getestet, variable Kausalzusammenhänge der charakterisierten Performance zu schätzen. In den Fallstudien im ersten Datensatz war die Bestimmung des primären Gebäudenutzungstyps dreimal treffender, die Bestimmung der Performance Klasse fast zweimal treffender und die Bestimmung des Betriebsverhaltenstyps über viermal treffender als für ein Basisvorgehen. Für die AMI Daten wurde die Charakterisierung der Standard Industrie Klasse um 27% verbessert, die Prognose der Erfolgsrate von Energiesparmassnahmen um 18% verbessert. Interviews mit Akteuren von mehreren Schulanlagen werden diskutiert bezüglich ihrer qualitativen Einblicke und bezüglich der Nützlichkeit der vorgestellten Ansätze im Kontext des Betriebs von Schulanlagen.