Yi Wang, Qixin Chen, Chongqing Kang

Smart Meter Data Analytics

Electricity Consumer Behavior Modeling, Aggregation, and Forecasting

September 23, 2019

Springer

To Our Alma Mater
–Tsinghua University

Foreword

Smart grid is a cyber-physical-social system where the power flow, data flow, and business flow are deeply coupled. Enlightened consumers facilitated by smart meters form the foundation of a smart grid. Countries around the world are in the midst of massive smart meter installations for consumers on the pathway towards grid digitalization and modernization. It enables the collection of extensive fine-grained smart meter data, which could be processed by data analytical techniques, especially now widely available machine learning techniques. Big data and machine learning terms are widely used nowadays. People from different industries try to apply advanced machine learning techniques to solve their own practical issues. The power and energy industry is no exception. Smart meter data analytics can be conducted to fully explore the value behind these data to improve the understanding of consumer behavior and enhance electric services such as demand response and energy management.

This book explores and discusses the applications of data analytical techniques to smart meter data. The contents of the book are divided into three parts. The first part (Chapters 1-2) provides a comprehensive review of recent developments of smart meter data analytics and proposes the concept of "electricity consumer behavior model". The second part (Chapters 3-5) studies the data analytical techniques for smart meter data management, such as data compression, bad data detection, data generation, etc. The third part (Chapters 6-12) conducts application-oriented research to depict the electricity consumer behavior model. This work includes electrical consumption pattern recognition, personalized tariff design for retailers, sociodemographic information identification, consumer aggregation, electrical load forecasting, etc. The prospects of future smart meter data analytics (Chapter 13) are also provided as the end of the book. The authors offer model formulations, novel algorithms, in-depth discussions, and detailed case studies in various chapters of this book.

One author of this book, Prof. Chongqing Kang, is a professional colleague. He is a distinguished scholar and pioneer in the power and energy area. He has done extensive work in the field of data analytics and load forecasting. This is a book worth reading; one will see how much insight can be gained from smart meter data alone. There are definitely broader qualitative understanding that can be gained from massive data collected in the realm of generation, transmission, distribution, and end-use of the smart grid.

September 2019

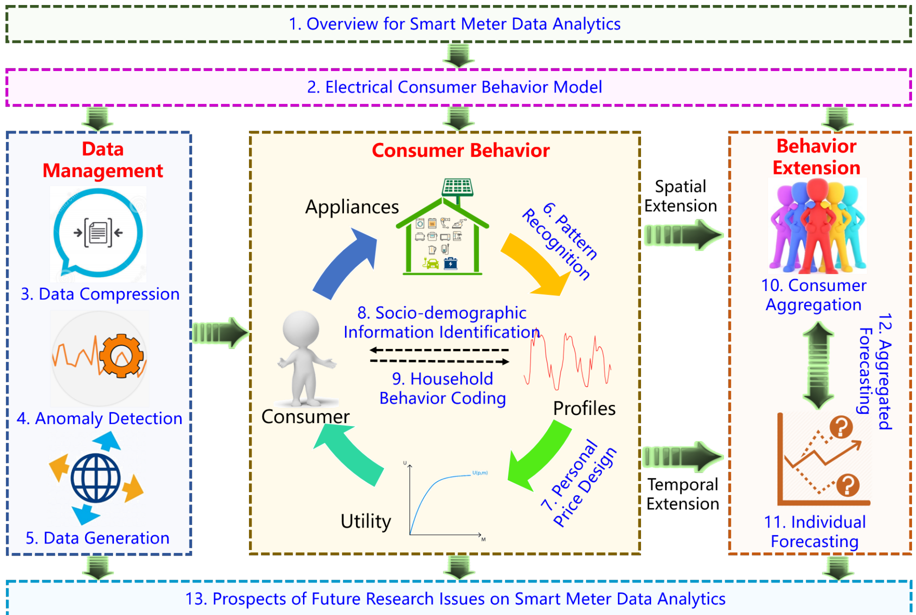
Prof. Saifur Rahman
Joseph Loring Professor and Founding Director
Advanced Research Institute at Virginia Tech
President of the IEEE Power and Energy Society

Preface

Decarbonization, decentralization, and digitalization (3D) are three pathways to future power and energy systems modernization. Most of the developments of the power and energy industry mainly focus on the generation and transmission sectors, while there is still a long way to go for distribution and demand sectors. Distribution systems in the electric power system have recently seen an important influx of exciting smart grid technologies such as distributed energy resources (DERs), multiple energy systems integration, control infrastructure, and data-gathering equipment. Increasing renewable energy integration and improve energy efficiency are two effective approaches for decarbonization. However, increasing penetration of renewable energy integration challenges the reliability, economy, and flexibility (REF) of the power and energy systems. A large number of DERs such as distributed photovoltaic (PV) and electric vehicles make the ADS more decentralized and complex. Broad interaction between consumers and systems can help provide flexibility to the power system and realize personalized consumer service. Meanwhile, information acquisition devices such as smart meters are gaining popularity, which enables an immense amount of fine-grained electricity consumption data to be collected. The "cyber-physical-social" deep coupling characteristic of the power system becomes more prominent. Breakthroughs are needed to analyze the behavior of electricity consumers.

Data analytics and machine learning techniques such as deep learning, transfer learning, graphical models, sparse representation, etc., have been greatly and considerably developed in recent years. It seems natural to figure out how to apply these state-of-the-art techniques to consumer behavior analysis and distribution system operation. However it is a predicament in the power industry that even though an increasing and huge number of smart meter data are collected and accessible to retailers and distribution system operators (DSOs), these data are not yet fully utilized for a better understanding of consumer behavior and an enhancement on the efficiency and sustainability of the power grid.

This book aims to make the best use of all of the data available to process and translate them into actual information and incorporate into consumer behavior modeling and distribution system operations. The research framework of the smart meter data analytics in this book can be summarized in the following figure.



This book consists of thirteen chapters. It begins with an overview of recent developments of smart meter data analytics and an introduction on the electricity consumer behavior model (ECBM). Since data management is the basis of further smart meter data analytics and its applications, three issues on data management, *i.e.*, data compression, anomaly detection, and data generation, are subsequently studied. The main components of electricity consumer behavior model include the consumer himself, appliances, load profiles, and the corresponding utility function. The following works try to model the relationships among these components and discover the inherent law within the behavior. Specific works include pattern recognition, personalized price design, socio-demographic information identification, and household behavior coding. On this basis, this book extends consumer behavior in both spatial and temporal scales. Works such as consumer aggregation, individual load forecasting, and aggregated load forecasting are introduced. Finally, prospects of future research issues on smart meter data analytics are provided.

To help readers have a better understanding of what we have done, we would like to make a simple review of the thirteen chapters in the following.

Chapter 1 conducts an application-oriented review of smart meter data analytics. Following the three stages of analytics, namely, descriptive, predictive and prescriptive analytics, we identify the key application areas as load analysis, load forecasting, and load management. We also review the techniques and methodologies adopted or developed to address each application.

Chapter 2 proposes the concept of ECBM and decomposes consumer behavior into five basic aspects from the sociological perspective: behavior subject, behavior environment, behavior means, behavior result, and behavior utility. On this basis, the research framework for ECBM is established.

Chapter 3 provides a highly efficient data compression technique to reduce the great burden on data transmission, storage, processing, application, etc. It applies the generalized extreme value distribution characteristic for household load data and then utilizes it to identify load features including load states and load events. Finally, a highly efficient lossy data compression format is designed to store key information of load features.

Chapter 4 applies two novel data mining techniques, the maximum information coefficient (MIC) and the clustering technique by fast search and find of density peaks (CFSFDP), to detect electricity abnormal consumption or thefts. On this basis, a framework of combining the advantages of the two techniques is further proposed to boost the detection accuracy.

Chapter 5 proposes a residential load profiles generation model based on the generative adversarial network (GAN). To consider the different typical load patterns of consumers, an advanced GAN-based on the auxiliary classifier GAN (ACGAN) is further to generate profiles under typical modes. The proposed model can generate realistic load profiles under different load patterns without loss of diversity.

Chapter 6 proposes a K-SVD-based sparse representation technique to decompose original load profiles into linear combinations of several partial usage patterns (PUPs), which allows the smart meter data to be compressed and hidden electricity consumption patterns to be extracted. Then, a linear support vector machine (SVM)-based method is used to classify the load profiles into two groups, residential customers and small and medium-sized enterprises (SMEs), based on the extracted patterns.

Chapter 7 studies a data-driven approach for personalized time-of-use (ToU) price design based on massive historical smart meter data. It can be formulated as a large-scale mixed-integer nonlinear programming (MINLP) problem, Through load profiling and linear transformation or approximation, the MINLP model is simplified into a mixed-integer linear programming (MILP) problem. In this way, various tariffs can be designed.

Chapter 8 investigates how much socio-demographic information can be inferred or revealed from fine-grained smart meter data. A deep convolutional neural network (CNN) first automatically extracts features from massive load profiles. Then SVM is applied to identify the characteristics of the consumers. Different socio-demographic characteristics show different identification accuracies.

Chapter 9 uses smart meter data to identify energy behavior indicators through a cross-domain feature selection and coding approach. The idea is to extract and connect customers' features from the energy domain and demography domain. Smart meter data are characterized by typical energy spectral patterns, whereas household information is encoded as the energy behavior indicator. The proposed approach offers a simple, transparent and effective alternative to a challenging cross-domain matching problem with massive smart meter data and energy behavior indicators.

Chapter 10 proposes an approach for clustering of electricity consumption behavior dynamics, where "dynamics" refer to transitions and relations between consumption behaviors, or rather consumption levels, in adjacent periods. To tackle the challenges of big data, the CFSFDP technique is integrated into a divide-and-conquer approach toward big data applications.

Chapter 11 offers a format of short-term probabilistic forecasting results in terms of quantiles, which can better describe the uncertainty of residual loads, and a deep-learning-based method, quantile long-short-term-memory (Q-LSTM), to implement probabilistic residual load forecasting. Experiments are conducted on an open dataset. Results show that the proposed method overrides traditional methods significantly in terms of pinball loss.

Chapter 12 proposes an ensemble method to forecast the aggregated load with sub-profiles where the multiple forecasts are produced by different groupings of sub-profiles. Different aggregated load forecasts can be obtained by varying the number of clusters. Finally, an optimal weighted ensemble approach is employed to combine these forecasts and provide the final forecasting result.

Chapter 13 discusses some research trends, such as big data issues, novel machine learning technologies, new business models, the transition of energy systems, and data privacy and security.

To summarize, this book provides various applications of smart meter data analytics for data management and electricity consumer behavior modeling. For each specific application, it either applies off-the-shelf machine learning methods to smart meter data and obtain novel results or makes some modification on traditional machine learning methods to make it more suitable for smart meter data. We hope this book can inspire readers to define new problems, apply novel methods, and obtain interesting results with massive smart meter data or even other monitoring data in the power systems.

Tsinghua University, Beijing
September 2019

Yi Wang
Qixin Chen
Chongqing Kang

Acknowledgements

This book made a summary of our research about smart meter data analytics achieved in recent years. These works were carried out in the Energy Intelligence Laboratory (EILAB), Department of Electrical Engineering, Tsinghua University, Beijing, China.

Many people contributed to this book in various ways. The authors are indebted to Prof. Daniel Kirschen from the University of Washington, Prof. Furong Li and Dr. Ran Li from the University of Bath, Dr. Tao Hong from the University of North Carolina at Charlotte, Dr. Ning Zhang, Dr. Xing Tong, Mr. Kedi Zheng, Mr. Yuxuan Gu, Mr. Dahua Gan, and Mr. Cheng Feng from Tsinghua University, who have contributed materials to this book.

We also thank Mr. Yuxiao Liu, Mr. Qingchun Hou, Mr. Haiyang Jiang, Mr. Yinxiao Li, Mr. Pei Yong, Mr. Jiawei Zhang, Mr. Xichen Fang, and Mr. Tian Xia at Tsinghua University for their assistance in pointing out typos and checking the whole book.

In addition, we acknowledge the innovative works contributed by others in this increasingly important area especially through IEEE Power & Energy Society Working Group on Load Aggregator and Distribution Market, and appreciate the staff at Springer for their assistance and help in the preparation of this book.

This book is supported in part by the National Key R&D Program of China (2016YFB0900100), and in part by the Major Smart Grid Joint Project of National Natural Science Foundation of China and State Grid (U1766212). The authors really appreciate their supports.

Yi Wang
Qixin Chen
Chongqing Kang

Contents