1-1-2017

Data-centric Situational Awareness and Management in Intelligent Power Systems

Xiaoxiao Dai
University of Denver

Data-centric Situational Awareness and Management in Intelligent Power Systems

---

A Dissertation

Presented to

the Faculty of the Daniel Felix Ritchie School of

Engineering and Computer Science

University of Denver

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

Xiaoxiao Dai

November 2017

Advisor: Jun Zhang

Author: Xiaoxiao Dai
Title: Data-centric Situational Awareness and Management in Intelligent Power Systems
Advisor: Jun Zhang
Degree Date: November 2017

Abstract

The rapid development of technology and society has made the current power system a much more complicated system than ever. The request for big data based situation awareness and management becomes urgent today. In this dissertation, to respond to the grand challenge, two data-centric power system situation awareness and management approaches are proposed to address the security problems in the transmission/distribution grids and social benefits augmentation problem at the distribution-customer lever, respectively.

To address the security problem in the transmission/distribution grids utilizing big data, the first approach provides a fault analysis solution based on characterization and analytics of the synchrophasor measurements. Specifically, the optimal synchrophasor measurement devices selection algorithm (OSMDSA) and matching pursuit decomposition (MPD) based spatial-temporal synchrophasor data characterization method was developed to reduce data volume while preserving comprehensive information for the big data analyses. And the weighted Granger causality (WGC) method was investigated to conduct fault impact causal analysis during system disturbance for fault localization. Numerical results and comparison with other methods demonstrate the effectiveness and robustness of this analytic approach.

As more social effects are becoming important considerations in power system management, the goal of situation awareness should be expanded to also include achievements in social benefits. The second approach investigates the concept and application of social energy upon the University of Denver campus grid to provide management improvement solutions for optimizing social cost. Social element – human working productivity cost, and economic element – electricity consumption cost, are both considered in the evaluation of overall social cost. Moreover, power system simulation, numerical experiments for smart building modeling, distribution level real-time pricing and social response to the pricing signals are studied for implementing the interactive artificial-physical management scheme.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation and Background

As the precondition for normal power system operations, security problems always receive the highest priority in power system operation and management. After the first wide-area blackout occurred in Northeast United State and Canada in 1965, condition monitoring of the power system became an important subject [1]. System status should be constantly monitored, that abnormal conditions should be detected and identified once it appears, or even before it happens, and system operators should be notified to take proper actions. The system conditions, *e.g.* voltage, current, frequency and power flow, are changing all the time to meet the generation and consumption balance, thus successful system monitoring requires adequate data, accurate analysis, efficient visualization tools and well trained operators to make wise decisions.

Although the importance of system condition monitoring has been addressed and a lot of efforts has been made, the development of human society keeps pushing its complexity to new levels. Aging of old facilities and infrastructures results in higher probability of equipment failure; rapid increase in electricity demand places

even greater burdens on the system and forces the power grids to expand; growth of renewable energy resources adds unstable elements to the grids; integration of large-scale distributed energy resources (DERs) introduces more uncertainties to the power system and is gradually changing the structure of the traditional power grids; what's more, concerns about environmental issues and rising of demand response are changing the way of power system operation and management.

Big or small blackouts started to appear across the world after 1965. After the 2003 Southeast U.S. and Canada blackout, the largest blackout in American history, the concept of "situation awareness" was officially put forward in the final report of task force [2]. And as a result of fast technology development, a vast amount of smart sensors, *e.g.* phasor measurement units (PMUs), frequency disturbance recorders (FDRs), smart meters, and sensors for building monitoring, have been widely installed to insure adequate data are recorded and provided for power system situation awareness. Not only the amount of available data is increasing rapidly, many more new types of data are also introduced into the power system, such as weather event data and health condition monitoring data, and following the trend of technology development, power system entered the Era of big data.

Coming of the Era of big data brings opportunities as well as challenges. Extremely large data sets guarantee that every moment and every aspect of the power system are monitored, but effective and efficient big data analyses become an issue. How to pick out real useful information from the data, how to increase the data processing speed, how to make best use of the data and how to make the data actionable are some of the major concerns of the researchers and operators. In the last few decades, researches on the data-driven/model-based approaches, *e.g.* machine learning and deep learning, proved that they can be a very promising solution to the big-data analysis. Compared with the computational approaches, data-centric system modeling can make better use of the measurement data, because it can dig

out more hidden information from the data, some of which are not even intuitively understandable. What's more, the data-centric modeling approaches have gradually exhibited the ability to work not only as an passive analytic tool, but also actively interact with the real system, providing advices while optimizing itself, which may make the system real "smart".

Extremely large scale data based situation awareness not only provides solutions for security problems in power systems. Upon secure and stable system operations, grid owners, utility companies and even users started to pursue higher goals, *e.g.* economic profits and social benefits. Demand side management becomes a very popular and important subject then, because its goal is to manage demands from the users' side to relieve the burden of peak time generation and transmission, which enlarges the benefits of both the grids and the users. Besides, installation rate of resident renewable generations and interests in electrical vehicles are climbing to respond to the environmental protection policies, which also brings the end users to the front of the power system managements. Involvement of end users brought in even more data with extremely high uncertainty and variability. User experiences, such as comfort level, health conditions and working efficiency, are critical feedbacks for intelligent power systems to improve their management strategies at the distribution-customer level.

In this dissertation, two data-centric situation awareness and management approaches are proposed to address security and social benefits augmentation problems in the intelligent power systems. Although occurrence of faults is much frequent at the distribution level, failures happened at the transmission level usually lead to severe consequences. Therefore, the first data-centric approach in this paper focuses on fault analysis algorithms for both transmission and distribution grids. And the second approach emphasizes the social benefits in the distribution-customer grids because most frequent interactions between end users and the grids emerge here.

1.2 Contribution

This dissertation represents two data-centric situation awareness and management approaches for solving power system security problem in the transmission and/or distribution grids and social benefits augmentation problem in the distribution - customer grids. Contributions of this work are listed below:

1. A two-layer synchrophasor data based approach is proposed for power system situational awareness in the transmission/distribution grids. Specifically, in the distribution grids, where the security monitoring system is not as mature as that in the transmission grids, and the topology information of which in some area is incomplete or totally missing, the only efficient way to detect and analyze faults is data-centric approaches.

    (a) The proposed approach characterized the PMU data in the spatial domain through secondary voltage control zones determination using a novel relative electrical distance calculation algorithm, which computes the electrical coupling ratio in dual-direction.

    (b) The optimal synchrophasor measurement devices selection algorithm (OSMDSA) is developed on the basis of optimal sunchrophasor measurement devices placement algorithm (OSMDP), which reduce the number of synchrophasor measurement devices while keeping the smart grid fully observable.

    (c) The data is further characterized in the temporal domain applying matching pursuit decomposition.

    (d) The weighted Granger causality is introduced for fault causal impact analysis.

2. A model-based parallel computing approach is proposed for distribution - customer level management to optimize overall social benefits, and the University of Denver campus grid is selected as the test bench to demonstrate the capability of the approach.

   (a) Working efficiency cost of building occupants together with energy consumption cost are considered in this paper as parts of the social cost to illustrate the concept of "social energy".

   (b) The human working efficiency model quantifies human comfort level, and associates it to monetary value.

   (c) Neural network is used to establish test building consumption models, both static and dynamic models are investigated.

   (d) Distribution locational marginal pricing algorithm is proposed in this paper to predict and reflect interactions between electricity users and the electricity suppliers.

   (e) Neural Network based iterative adaptive dynamic programming algorithm is introduced for solving optimization problems of the complex highly nonlinear system.

1.3 Dissertation Outline

Arrangement for the rest of the dissertation is given as the following.

Chapter 2 first reviews related works on PMU development and power grid fault detection techniques. Then a review of the current demand-side management schemes and technologies is given.

Chapter 3 depicted the synchrophasor data based situational awareness approach in transmission grids. Specifically, the optimal synchrophasor measurement devices selection (OSMDSA) algorithm, hidden Markov models based fault detection

method and weighted Granger causality based fault impact analysis are presented in detail.

Chapter 4 describes a socio-technical power system under the concept of social energy. University of Denver's campus grid was used as a case study, and its HVAC management strategy is studied aiming at maximizing social energy utilization. To form the socio-technical system of the campus grid, neural network based consumption models, human working efficiency quantification, distribution locational marginal pricing are introduced. For solving optimal HVAC strategy, two cases, where the system is assumed to be static and dynamic respectively, and distributed iterative adaptive programming algorithm and Neural Network based iterative adaptive dynamic programming are introduced and applied in the two case, respectively.

At last, Chapter 5 concludes all the works done for this dissertation and proposed possible future directions of the current research.

Chapter 2

Literature Review

2.1 Current situation of modern power grid security monitoring

With the evolution of energy and power technology, the emerging Smart Grid (SG), featured by security monitoring, distributed renewable energy generation and demand-response control, is facing a new challenge in extremely large data sets, namely the Big Data challenge [3, 4]. The data sources of a future electrical SG, are heterogenous, including smart metering devices such as phasor measurement units (PMUs) and frequency disturbance recorders (FDRs), automated revenue metering at the electricity market side, synchrophasor measurement data for situational awareness, SG component state and operation control data, as well as large data sets from sources other than SG measurements, such as geographic information system data. For example, in the State Grid Corporation of China, more than 2500 PMUs are equipped at the end of 2013, and the number will increase in the future [5]. This means that the widely located synchrophasor measurement devices can provide abundant data, and the analysis and information extraction from the big data becomes an imperative issue for SG monitoring, security and control [6–8].

Based on the big data collected by the synchrophasor measurement devices, many big data based computational approaches are applied in smart grid to address different problems, such as smart grid monitoring, distributed control, renewable energy forecasting, etc. [9, 10]. In [11], a cloud-based software platform is designed for big data analysis in smart grid. A model for utilizing cloud computation is used for smart grid big data management in [12]. However, these approaches bring up the question of how to effectively and efficiently characterize and manage the power system big data, and therefore, it is critical and urgent to develop a characterization approach for this purpose. In the spatial domain, it is proposed to use the optimal synchrophasor measurement devices placement (OSMDP) to reduce the number of synchrophasor measurement devices while keeping the smart grid fully observable [13,14]. In temporal domain, the matching pursuit decomposition (MPD) with Gaussian atom dictionary is used to characterize the signals [15–17]. Based on these technical methods, in this paper, an effective and highly accurate data-driven solution of fault analysis is proposed to respond to the Big Data challenge in characterization and analytics of the synchrophasor measurement system data.

In this paper, we propose a framework for characterizing and managing the data generated by the synchrophasor measurement devices in the spatial-temporal domain. Specifically, in the spatial domain, the secondary voltage control (SVC) is used to divide the SG into a number of SVC zones. Within each SVC zone, a subset of synchrophasor measurement devices on their corresponding buses are selected based on an optimal synchrophasor measurement devices selection algorithm (OSMDSA), which is based on OSMDP. In [18, 19], the SVC is analyzed, and the determination algorithms for the SVC zones and the pilot buses are studied. In [20, 21], different algorithms are studied on OSMDP, and different test scenarios are considered and compared. Based on SVC and OSMDP, the proposed approach provides a hierarchical and dynamic spatial characterization of the synchrophasor

measurements. In the time domain we propose to use the MPD with Gaussian atom dictionary for fully characterizing the signals [15, 16].

In the last few decades, research of data-driven approaches and machine learning, e.g., the support vector machine (SVM) and artificial neural network (ANN), indicates that there is a promising way to detect and diagnose SG faults [22, 23]. Using the generated spatial-temporal characteristics, various hidden Markov models (HMMs) are trained for detection and identification of SG faults to achieve situational awareness. To identify and locate the major impact bus in the presence of a fault, Granger causality [24] is introduced to study the fault impact causality analysis among all the buses in the SG.

2.2 Demand-side management schemes and techniques

Ever since the term "demand side management (DSM)" was proposed in 1979 and was firstly introduced to the public by Electric Power Research Institute in the 1980s [25, 26], nearly four decades has passed. The rapidly increased power demand, newly integrated energy forms new patterns of energy market operations, and policy and social requirements all push the DSM techniques to improve. In the meanwhile, breakthroughs in communication technology, enhancements in computation power, new forms of energy supply and storage, and the expanding involvement of consumers into energy management rendered DSM new possibilities and variations.

2.2.1 Demand-side management in residential, industrial and commercial sectors

Residential houses, which nowadays are capable of integrating PV panels, various smart appliances and electrical vehicles, seem to offer the highest flexibility for DSM. Besides, the household's flexible living schedule, or the different living patterns of households (regarding DSM for a community other than for just a single house), make DSM for residential buildings even more elastic.

Both task-scheduling and energy-managing schemes are applied in residential DSM [27]. Task-scheduling schemes focus on the activation/deactivation time of schedulable appliances (e.g. dish washers, washing machine, dryer, rice cookers, and electrical vehicles' battery charging) and/or distributed energy resources (DERs), in response to market signals (price changes or incentives) or availability of renewable generations. While energy-managing schemes deal with the operation of non-schedulable devices, such as HVAC systems, heat pumps, and refrigerators, aiming at reduce their consumption or enhance their energy efficiency.

In the case study of [28], the optimal scheduling problem of five DERs, plug-in hybrid vehicle, space heater, storage water heater, pool pump, and the PV system in a smart home environment was studied. The co-evolutionary version of particle swarm optimization (CPSO) was applied to solve the scheduling problem in order to maximum net benefit while keeping the living condition in a comfortable range. In [29], driven by the dynamic electricity pricing system adopted in Jeju Island, the authors implemented genetic algorithm for scheduling 15 appliances in a smart home, in pursue of lower electricity cost. While in [30] and [31], load-shifting problem of schedulable loads and consumption reduction problem for non-schedulable loads for multiple users/buildings are combined and studied.

Despite the huge potential of residential DSM, the largest obstacle in front of its real world applications might be the willingness of the households to participate in such programs. Although some utility companies offer incentives or put penalties to encourage or force their customers to participate in DSM programs, the fear of private information reveal, relatively low profitability comparing to the installation and maintenance costs of smart meters and devices, concerns of uneven incentives, or even no elasticity for load shifting (for those whose energy consumption only covers basic living needs) can prevent the households to support and join such programs.

DSM programms have much larger market in the non-residential sectors. According to the latest energy review by the U.S. Energy Information Administration [32], non-residential sector, which consists of transportation, industrial and commercial, consumed 79% of the overall energy in 2016. Since the term "demand" in transportation is not in the sense of "energy", demand management in transportation refers to different concepts and goals [33, 34], thus we only talk about DSM in the industrial and commercial sectors here. Compared to single household, industry and commercial owners should be much more eager and positive to the DSM programs in order to reduce energy costs and avoid penalty for over-consumption at peak periods [35]. Besides, the larger scale of these sectors also gives opportunities for DSM to make "louder sound".

For industries, DSM tasks are more specific and specialized, for they need to deal with certain types of devices and production processes, and very specific requirements need to be meet to ensure safe and successful production. [36] conducted an empirical investigation on the potential, strategies and barriers of DSM applications in the refrigerated systems in German. And in [37], the researchers evaluated the feasibility and efficiency of DSM in cement-making industry by discussing the equipments and processes in cement production, operational limitations of cement plants, and the amounts and sources of energy used in the cement-making processes.

For commercial customers, DSM are mostly applied on buildings, including offices, academic buildings, laboratories, gyms and arenas and so on [38–40]. Unlike residential houses which reflects very high flexibility, nor industrial sites, operation and energy usage patterns of which follow strict rules, commercial buildings are something in between. DSM for commercial customers generally doesn't have high temporal flexibility, for example during working hours (for offices, classrooms and laboratories) or events (for arenas and stadiums), certain equipments and facilities have to be on and work for certain hours. But on the other hand, buildings of the

same type may have similar consumption patterns [27]. DSM of commercial buildings are generally realized through control of HVAC systems, lighting systems and operation modifications of heavy load equipments [41–44].

2.2.2 Integrations of human factors in demand-side management development

In early studies of DSM, the goals/objects were simply the reduction of peak-hour loads or reduction of energy costs. Later, other concerns were taken into account, for instance the emission costs, startup, operation and maintenance costs [45–47]. Number of DSM schemes with multiple objects is increasing and the scenarios and algorithms are becoming more complicated. The study in [48] tries to minimize both the operational cost and load power for residential customers, [49] aims at maximizing the expected market profits for the retailer and minimizing electricity cost for consumer at the same time, and [31] tries to achieve minimum power consumption while ensuring maximum social welfare for residential households.

More recently, the researchers started to notice the human factors in DSM, either the effect of users' behavior on consumption patterns, or the importance of users' experience with the DSM programs. [43] research investigated the energy consumption profile in a multi-purpose academic building with respect to the occupants behavior, and indicates strong relationship between consumption and occupants behavior (e.g. number of occupants, duration of staying in the buildings). [42, 50] investigate electricity consumption model of office buildings with respect to occupants' behavior. [31, 51–53] implemented users' comfort/discomfort/satisfaction level to the DSM programs they developed. And [54, 55] input user preferred temperature ranges as system constraints.

Back into the three sectors we discussed above, effects of human behavior on the consumption pattern in residential houses could be very complicated to model,

since the living habits, preference and criteria towards life quality are quite different from person to person. However, households may exhibit higher tolerance to the DSM control schemes. Using HVAC control as an example, as long as the indoor temperature is in the comfort range, the households won't feel huge difference of the temperature. While this is quite different in the commercial sector, for temperature changes directly affect the workers'/users' working performance [56, 57], which could be in the top priorities of business owners, universities or research institutes. Behavior effects of workers in the industrial sector are in another sense, for their activities directly affect the production processes. And these kinds of behavior effects are the most difficult ones to model and evaluate, but play very important roles in industry. For better understanding and implementation of human factors in the DSM schemes, complicated behaviors modeling techniques and investigations on behavior quantification have to be improved and conducted.

2.2.3 Techniques and algorithms used in current demand-side management programs

Traditional optimization algorithms have been widely used in DSM. For example, linear programming [48, 52, 58], reinforcement learning [59], Lagrangian algorithm [60], convex optimization [31,61], and particle swarm algorithm [28]. As the objects of DSM programs becoming compound, the algorithms are becoming more complex, that only sub-optimal solutions can be found, or computational complexity would be too high to find optimal solutions. Therefore, heuristic optimization were introduced [44]. What's more, utilization of distributed algorithms and game theories [30,49,62] is also increasing.

Chapter 3

Synchrophasor Data Based Situational Awareness in Transmission/distribution Grids

3.1 Formulation and Summary of the Proposed Approach

Based on the background of increased PMU utilization in SGs [5], in this chapter, it is assumed that the synchrophasor measurement devices are located on every bus of the SGs, and the voltage measurements $\mathbf{S}$ is denoted as $\{s_1(t), s_2(t), s_3(t), \cdots, s_n(t)\}$, which are collected by the widely located synchrophasor measurement devices, where t denotes time, and $s_i(t)$ is the voltage signal from bus i.

Considering the computation complexity of the proposed approach, the system operation state of SG is divided into two states: normal state Λ^0, and fault state $\mathbf{\Lambda}^\phi$, which contains M_1 fault types $\Lambda^{(m)}$, $m = 1, 2, 3, \cdots, M_1$. As shown in Fig. 3.1, the proposed approach can be divided into two data analytic layers. The first layer is used to determine the SVC zones $\mathbf{\Omega}_{i_1}$ and pilot buses $\mathbf{\Psi}$ in the normal state of SG,

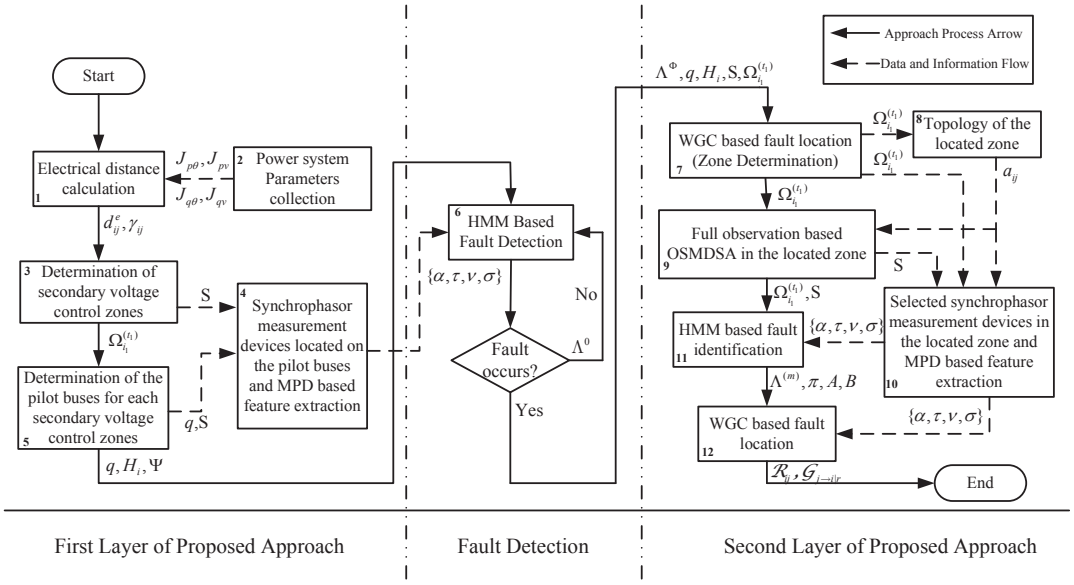


Figure 3.1: The flowchart of proposed approach.

where $i_1 \in \{1, 2, 3, \cdots, k\}$. The second layer is used to identify the fault and locate the fault in the fault located SVC zone. Between the two layers, a fault detection part based on HMM is designed as a trigger to start the second layer.

As shown in Fig. 3.1, in the first layer, the system parameters are collected for relative electrical distance (RED) calculation (Block 1 and 2). Then, with the k-means clustering algorithm, the SVC zones $\mathbf{\Omega}_{i_1}$ are determined, and pilot buses $\mathbf{\Psi}$ of the given SVC zones are also determined (Block 3 and 5). In Block 4, the collected voltage signals from the pilot buses are processed using MPD for feature extraction and noise suppression.

Between the first layer and second layer, an HMM based detector is designed as a trigger for the second layer. Specifically, in Block 6, two trained HMMs Λ^0 and $\mathbf{\Lambda}^\phi$ are used to detect if the SG is in abnormal states. If the fault state is detected, the second layer of the proposed approach is triggered for fault identification and location.

In the second layer, using the voltage measurements from the pilot buses $\mathbf{\Psi}$, the fault can be located in a SVC zone with WGC (Block 7). With the topology information of the located SVC zone, the OSMDSA is used to select the synchrophasor measurement devices in the located SVC zone, which ensures full observability (Block 8 and 9). In Block 10, the MPD is used to extract features and suppress noise from the voltage signals. In Block 11, the fault is identified with different fault modelled HMMs. In Block 12, the fault is located using WGC.

In some scenarios, a fault occurs on the edge of several adjacent SVC zones. In these conditions, aiming to locate the fault, the proposed approach can selected them according to WGC determination (Block 7). In Block 9 and 10, the OSMDSA can dynamically treat the selected SVC zones as a larger combined SVC zone, and compute the optimal synchrophasor measurement devices for fault identification and location.

3.2 Spatial-Temperal Synchrophasor Data Characterization

3.2.1 Determination of SVC Zones and Pilot Buses

The approach for determining the SVC zones and pilot buses is described as following.

Relative Electrical Distance

There are several methods to define and calculate the RED for the SVC zones and pilot buses [19, 63–65]. In this chapter, the proposed approach is focusing on wide area monitoring for SG with big data background. The RED is based on the linearized model of the power flow equation as following

$$\begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix} = \begin{bmatrix} J_{p\theta} & J_{pv} \\ J_{q\theta} & J_{qv} \end{bmatrix} \times \begin{bmatrix} \Delta\theta \\ \Delta V \end{bmatrix}. \tag{3.2.1}$$

where ΔP and ΔQ are two vectors denoting the variances of active power and reactive power, respectively; $\Delta\theta$ and ΔV are vectors denoting the variances of voltage angle and voltage magnitude, respectively; $J_{p\theta}$, J_{pv}, $J_{q\theta}$, and J_{qv} are the sub-matrices of Jacobian matrix. Equation (3.2.1) is the power flow equation from the Newton-Raphson method, which indicates the relationship between the variance of active and reactive power, and the variance of voltage angle and voltage magnitude [66].

The voltage is strongly coupled with reactive power and weakly coupled with active power. Generally speaking, the reactive power plays an important role in voltage control and voltage stability analysis [64, 66]. In this chapter, the injection active power is set as constant, which means $\Delta P = 0$, then the sensitive between reactive power and voltage can be computed as following

$$SE = \Delta V/\Delta Q = (J_{qv} - J_{q\theta}J_{p\theta}^{-1}J_{pv})^{-1}. \tag{3.2.2}$$

The voltage sensitivity between bus i and j can be derived as

$$\gamma_{ij} = \frac{\Delta V_i}{\Delta V_j} = (\frac{\Delta V_i}{\Delta Q_j})/(\frac{\Delta V_j}{\Delta Q_j}) = \frac{SE_{ij}}{SE_{jj}} \tag{3.2.3}$$

where ΔV_i and ΔV_j are the voltage variances at bus i and j, respectively; ΔQ_j is the variance of reactive power at bus j. The electrical coupling ratio between bus i and j is defined as

$$d_{ij}^c \triangleq \sqrt{\gamma_{ij}^2 + \gamma_{ji}^2} \tag{3.2.4}$$

The electrical coupling ratio d_{ij}^c denotes the relationship of voltage coupling between bus i and j. If there is strong voltage coupling between bus i and j, the electrical coupling ratio d_{ij}^c is large. $d_{ij}^e \triangleq 1/d_{ij}^c$ represents the RED between bus i and j. If there is strong voltage coupling between bus i and j, the RED d_{ij}^e is small. Then the

RED matrix $\mathbf{D}^e$ can be built, where $\mathbf{D}^e \in \mathbf{R}^{n\times n}$, and d^e_{ij} is an element at row i, and column j. Furthermore, compared with the traditional single-direction electrical coupling ratio in [67, 68], the proposed approach computes the electrical coupling ratio in dual-direction as equation (3.2.4), which provides a more comprehensive measurement about the relative electrical distance between the two buses.

Determination of SVC zones

According to the above defined equations, the problem of SVC zones determination can be transferred to a clustering problem. In a given SG, the buses with small REDs are clustered into the same SVC zone, the buses with large REDs are clustered into different SVC zones. Through the algorithms such as genetic algorithm and annealing algorithm, the buses can be clustered into several SVC zones [63, 64]. In this chapter, the k-means algorithm is implemented to cluster the buses into different clusters.

First, an initial set of k means $\{\eta_1, \eta_2, \eta_3, \cdots, \eta_k\}$ of k corresponding clusters is given. Then, each bus in the smart grid is assigned to the cluster which yields the least distance between its mean to the element among all clusters as following

$$\mathbf{\Omega}^{(t_1)}_{i_1} = \{\xi_i : \boldsymbol{d}(\xi_i, \eta^{(t_1)}_{i_1}) \leqslant \boldsymbol{d}(\xi_i, \eta^{(t_1)}_{j_1}) \ \forall j_1, 1 \leqslant j_1 \leqslant k\}. \tag{3.2.5}$$

where t_1 is the iteration number; $i_1 \in \{1, 2, 3, \cdots, k\}$; $\xi_i = [d^e_{i1}\ d^e_{i2}\ ...\ d^e_{in}]$ is the ith row in $\mathbf{D}^e$, which indicates the RED from bus i to other buses; and $\boldsymbol{d}(\cdot, \cdot)$ represents the Euclidean distance function in the RED matrix $\mathbf{D}^e$. The new means of each cluster is updated as in (3.2.6) and the data set is re-clustered until the means converge [69, 70].

$$\eta^{(t+1)}_{i_1} = \frac{1}{|\mathbf{\Omega}^{(t)}_{i_1}|} \sum_{\xi_{j_2} \in \mathbf{\Omega}^{(t)}_{i_1}} \xi_{j_2}. \tag{3.2.6}$$

Determination of Pilot Bus

In a given SVC zone, a pilot bus is chosen to represent the voltage variances of the zone. Based on the discussion above, this condition requires the REDs between the pilot bus and the other buses are short. Let Ω_{i_1} be the bus set of SVC zone i_1, the total RED between bus i and other buses in the given SVC zone Ω_{i_1} is calculated as

$$H_i = \sum_{j \in \Omega_{i_1}} d_{ij}^e. \tag{3.2.7}$$

where j is the bus index in SVC zone Ω_{i_1}. Based on (3.2.7), the pilot bus q can be determined as following

$$q = \arg \min_i (H_i). \tag{3.2.8}$$

In some SVC zones, it is possible that multiple pilot buses are selected according to (3.2.7) and (3.2.8).

3.2.2 Optimal Synchrophasor Measurement Devices Selection Algorithm

Basic Optimal Synchrophasor Measurement Devices Selection Algorithm

The objectives of OSMDSA are to minimize the number of synchrophasor measurement devices, while to ensure full system observability. This means, in an observable SG, each bus must be observed at least once

$$f_i = \sum_{j=1}^{n} u_j \ \geqslant 1, \ i = 1, 2, 3, ..., n \tag{3.2.9}$$

where f_i refers to the number of times that the ith bus is observed through synchrophasor measurements, u_j is defined as a binary decision variable as

$$u_j = \begin{cases} 1 & \text{A synchrophasor measurement device is selected at bus } j \\ 0 & \text{otherwise} \end{cases} \tag{3.2.10}$$

Based on [71, 72], the concept of topological observability is adopted and the following rules are applied for optimal synchrophasor measurement devices selection.

- If the voltage phasor and current phasor at one end of a branch are known, the voltage phasor at other end of that branch can be obtained using Ohm's law.
- If the voltage phasors at both ends of a branch are known, the current phasor through this branch can be calculated.

The measurements such as bus voltage phasors and branch current phasors, directly obtained from the synchrophasor measurement devices, are referred to as direct measurements. Measurements derived by employing the above two rules are referred to as indirect measurements, or pseudo measurements. In an observable network, each and every bus must be observed at least once by using direct or indirect measurement. Then, based on (3.2.9) and (3.2.10) the problem can be formulated as

$$\min_{u_i} F = \sum_{i=1}^{n} c_i u_i \tag{3.2.11}$$

subject to constraints:

$$f_i = \sum_{j=1}^{n} a_{i,j} u_j \ \geqslant 1, \ i = 1, 2, 3, ..., n \tag{3.2.12}$$

where a_{ij} is the (i, j)th entry of system connectivity matrix defined as

$$a_{ij} = \begin{cases} 1 & \text{if } i = j \text{ or if } i \text{ and } j \text{ are connected} \\ 0 & \text{otherwise} \end{cases} \tag{3.2.13}$$

c_i is the cost of selecting a synchrophasor measurement device at bus i, which can be the economic cost, the communication cost, computation cost, or the combination of the above cost. In our problem, the cost of a synchrophasor measurement device selection at each bus is assumed to be 1 per unit.

The basic OSMDSA algorithm can ensure system observability when the system is in the normal operation states. In addition, it can also ensure system observability in the presence of certain types of faults, which do not change the topology of the SGs.

Optimal Synchrophasor Measurement Devices Selection Algorithm Considering Single Line Outage

Certain types of system faults, such as transmission line outage can change grid topology, and the OSMDSA approach described in (3.2.9)-(3.2.13) can then be extended to more robust derivative approaches, which are resilient to be these faults. In this chapter, we investigate the scenario where a single line outage occurs in the power system, and the resulting OSMDSA approach is able to provide data characterization with full system observability. It is noted that, other faulty scenarios, such as single synchrophasor measurement device disable, controlled islanding, etc, can also result in different spatial data characterizations.

Transmission line outage is a typical fault in SG [21, 66]. Single line outage (SLO) may cause loss of observability for one of its terminal buses, which would otherwise be observable using the current phasor of that line. For a power network with M lines. The constraints of SLO can be presented as

$$f_i = \sum_{j=1}^{n} a_{i,j}^{l} u_j \ \geqslant 1, \ \forall i, \forall l, \tag{3.2.14}$$

where $l \leqslant M$, the parameter $a_{i,j}^{l} = 0$ if the lth outage line is used to connect buses i and j, and $a_{i,j}^{l} = a_{i,j}$ otherwise.

In summary, jointly using the constraint (3.2.14) and the objective function (3.2.11), the OSMDSA-SLO approach is formulated, which provides the optimal data characterization for this scenario.

3.2.3 Time-Frequency Characterization of Synchrophasor Measurements

The MPD, proposed in [15,16], is an effective time-frequency analysis algorithm. In this chapter, MPD with Gaussian atom dictionary is used to characterize synchrophasor data in the temporal-domain and extract the signal features represented by the amplitude, time-shift, frequency-shift and variance of the Gaussian atoms.

Matching Pursuit Decomposition for Feature Extraction

On bus i, a continuous voltage signal $s_i(t)$ can be decomposed as a weighted summation of the Gaussian atoms as

$$s_i(t) = \sum_{p=1}^{\infty} \alpha_p \, g_p(t) \,, \tag{3.2.15}$$

where α_p is the amplitude coefficient for the Gaussian atom $g_p(t)$, $g_p(t)$ is selected from a given Gaussian atom dictionary, and p is the index. To prove the completeness of the expression (3.2.15), it is provided the following lemma.

Lemma 1: The voltage signal $s_i(t)$ can be represented using decomposition with finite iterations and a remainder $r_N(t)$ as

$$s_i(t) = \sum_{p=1}^{N-1} \alpha_p \, g_p(t) + r_N(t) \,, \tag{3.2.16}$$

with $||r_N(t)||$ decaying exponentially in a finite dimensional space.

Proof: With the notations in (3.2.16), the MPD algorithm is described as follows. Let $r_1(t) = s_i(t)$, and the atoms $g_p(t)$ is selected from a very redundant

Gaussian atom dictionary $\mathcal{D}$ as the one that has the maximum magnitude of the projection on $r_p(t)$, $p = 1, 2, 3 \cdots P$. Specifically $g_p(t)$ is obtained using

$$g_p(t) = \arg\max_{g^{(\mathbf{e})}(t) \in \mathcal{D}} \Big| \int_{-\infty}^{\infty} r_p(t)\, g^{(\mathbf{e})}(t) dt \Big| \,, \tag{3.2.17}$$

where $\mathbf{e} = \{\tau, \nu, \sigma\}$ are the time-shifting, frequency-shifting and related shape parameter for the Gaussian atoms. After $g_p(t)$ is obtained, the corresponding coefficient α_p is calculated as

$$\alpha_p = < r_p(t)\,, g_p(t) > = \int_{-\infty}^{\infty} r_p(t)\, g_p^*(t) dt \,, \tag{3.2.18}$$

where $g_p^*(t)$ is the complex conjugate of $g_p(t)$. (3.2.16) yields an energy conservation equation

$$||s_i(t)||^2 = \sum_{p=1}^{N-1} ||\alpha_p\, g_p(t)||^2 + ||r_N(t)||^2 \,. \tag{3.2.19}$$

The decay of $||r_N(t)||$ depends upon the correlation between the residues and the dictionary elements. It is defined that the correlation ratio of a function $s_i(t)$ with respect to $\mathcal{D}$ as

$$\varrho(s_i(t)) = \sup_{g^{(\mathbf{e})}(t) \in \mathcal{D}} \frac{< s_i(t), g^{(\mathbf{e})} >}{||s_i(t)||}. \tag{3.2.20}$$

And it can be proved that [15]

$$\kappa(\varrho) = \inf \varrho(s_i(t)) > 0. \tag{3.2.21}$$

Based on (3.2.20) and (3.2.21), in [15, 16] it is proved that the $||r_N(t)||$ decays exponentially in a finite dimensional space

$$||r_N(t)|| \leqslant ||s_i(t)||(1 - \iota^2 \kappa^2(\varrho))^{N/2} \,, \tag{3.2.22}$$

where ι is an optimality factor with $0 < \iota \leqslant 1$, and satisfies

$$| < r_N(t), g_N(t) > | \geqslant \ \iota \varrho(r_N(t)) ||r_N(t)|| \,. \tag{3.2.23}$$

Based on (3.2.19) and (3.2.22), the original signal $s_i(t)$ can be decomposed with limited residual in finite Gaussian atom dictionary with iteration N, and the fact is given by

$$||s_i(t)||^2 - \sum_{p=1}^{N-1} ||\alpha_p \, g_p(t)||^2 \leqslant \epsilon ||s_i(t)||^2 \,, \tag{3.2.24}$$

where $\epsilon = (1 - \iota^2 \kappa^2(\varrho))^{N/2}$ is a desired precision value that decays exponentially. It is able to conclude that expression (3.2.16) is complete in the sense of exponentially decaying residual energy.

Therefore, the MPD analysis can decompose a voltage signal with a controllable residual signal power, and usually, if a very redundant Gaussian atom dictionary is used, after a small number of decomposition iterations, the extracted signal features can be used to synthesize the original signal with very high fidelity. The voltage signal can be represented with a group of parameters $\{\alpha, \tau, \nu, \sigma\}$. As a result, it is expected that this approach can fully characterize the synchrophasor measurements in temporal domain.

An Example on Transmission Line Grounding in IEEE 39-Bus System

A transmission line ground fault in the IEEE 39-bus system is illustrated with MPD analysis in Fig. 3.2. The simulation duration is 4 s with a transmission line grounding fault occurring between bus 5 and bus 6 for 10 cycles. The voltage signal is collected from bus 4. The MPD uses 1,800,000 Gaussian atoms and 30 iterations. In this case, in order to test the robustness of the proposed MPD, the test signal-to-noise ratio (SNR) is 25 dB, which is much lower than the typical noise (72 dB) in synchrophasor measurement devices [73, 74].

As shown in Fig. 3.2(a), the green curve is the original voltage signal of transmission line grounding, which illustrates the large voltage deviation caused by fault from 2000 ms to 2200 ms. In Fig. 3.2(b), the red curve is the original voltage signal

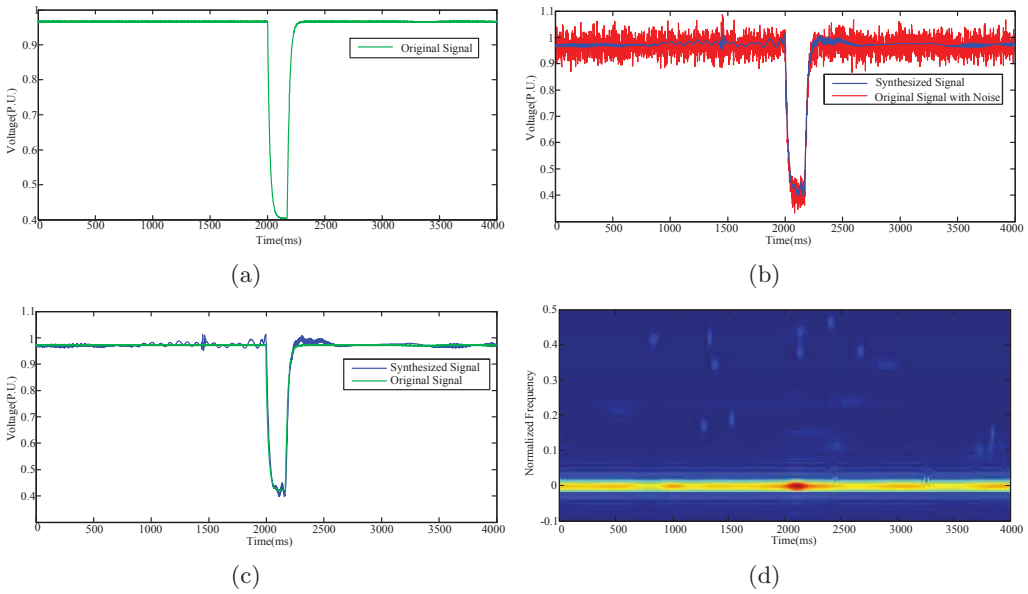


Figure 3.2: (a) The original signal of voltage drop curve, (b) the original signal with noise and the synthesis curve using MPD, (c) the original signal and synthesis curve using MPD, (b) the time-frequency analysis of the synthesis curve using MPD.

with noise and the blue curve is the synthesis voltage signal, which is extracted from the red curve using the composition of selected Gaussian atoms. The synthesized signal suppresses the noise of the red curve, and recovers the original voltage signal, especially for the large voltage deviation caused by the fault between 2000 ms to 2200 ms. In Fig. 3.2(c), this is the comparison between the original signal and synthesis signal. Although there are small voltage deviations between the original and the synthesized voltage signals, for example a small voltage deviation occurs about 1500 ms, the energy of the residual signal takes only 0.03% of the original signal. And the ratio of residual signal energy can be computed as following:

$$r_{reen} = \frac{||r'_N(t)||^2}{||r'_1(t)||^2}, \tag{3.2.25}$$

where $r'_1(t)$ is the original signal without noise, $r'_N(t)$ is the residual signal between the original signal $r'_1(t)$ and the synthesized signal, and $r_1(t)$ is the original signal

with noise. As Fig. 3.2(d) illustrates, although there are a lot of noise, the original voltage signal can be recovered using several Gaussian atoms with a controllable residual energy, which are used to represent the feature of the original signal. This indicates the proposed MPD can extract the feature of the voltage signal and suppress the noise.

3.3 Data Analysis for Fault Detection, Identification and Location

3.3.1 Hidden Markov Models for Fault Detection and Identification

Using the spatial characterization obtained by OSMDSA, and the temporal characterization by MPD, the SG situational awareness can be achieved. In this section, a data-driven SG fault detection and identification approach are investigated to validate and evaluate the proposed data characterization approach.

The detailed process of fault detection is depicted as following. Different HMMs are trained to detect and identify different types of faults in SG for situational awareness. In this chapter, five typical faults in SGs are studied, which are generator ground, load loss, generator outage, single-phase transmission line outage, and three-phase transmission line outage. As shown in Block 6 in Fig. 3.1, two HMMs, Λ^0 and $\mathbf{\Lambda}^\phi$ are trained and then used to detect normal and abnormal operating conditions, respectively. Specifically, Λ^0 is the HMM for normal conditions, which is trained by the collected data in normal operation conditions. $\mathbf{\Lambda}^\phi$ is the HMM for abnormal conditions, which is trained by the collected data in abnormal operation conditions as mentioned before.

Through training, for the normal operation conditions, the parameters $\lambda^0 = \{\pi^0, A^0, B^0\}$ are computed to represent the initial states distribution vector, the hidden state transition matrix, and the state-dependent observation matrix, respec-

tively. Using the Baum-Welch algorithm [75], the maximum-likelihood estimate for λ is given by

$$\lambda_{ML} = \arg \max_{\lambda} \log P(\boldsymbol{\beta}|\lambda, \Lambda)\,, \tag{3.3.1}$$

where $\boldsymbol{\beta}$ is the observation symbol sequence obtained from MPD, λ is the parameter set of HMM Λ. For the abnormal operations, the parameters $\lambda^{\phi}=\{\pi^{\phi}, A^{\phi}, B^{\phi}\}$ are computed similarly. Then, using testing data, a confusion matrix is generated from the detection results to evaluate the detection performance of the proposed approach. A detection result using IEEE 39-bus system is shown in Table 3.1, which indicates the high performance of the normal/abnormal detection.

Table 3.1: Confusion Matrix of Fault Detection Between Normal and Abnormal conditions in IEEE 39-bus System

	Normal	Abnormal
Normal	**100%**(280/280)	0(0/280)
Abnormal	0(0/280)	**100%**(280/280)

In fault identification, for the mth fault type, an HMM $\Lambda^{(m)}$ is trained to identify the fault type m after the fault occurs. In $\Lambda^{(m)}$, the parameters $\lambda^{(m)}=\{\pi^{(m)}, A^{(m)}, B^{(m)}\}$ represent the initial state distribution vector, the hidden state transition matrix, and the state-dependent observation density matrix, respectively [75].

As the faults of the generator ground, load loss, and generator outage do not change the topology of the SGs, the basic OSMDSA is used for spatial characterization when we test these hypothesis in fault identification. The single phase transmission line outage and three-phase transmission line outage can change the topology of SGs. As a result, when we test the hypothesis correponding to these two fault scenarios, the SG can be observed by the results of OSMDSA with SLO.

3.3.2 Weighted Granger Causality Based Fault Causal Impact Analysis

Using the Granger causal analysis, causal relationships between any pair of the buses in the system can be computed, thus the effect of a certain bus on all the other buses can be generated and presented. Let $\mathbf{s}_i = [s_i(1)\ s_i(2)\ \cdots s_i(l_w)\ \cdots\ s_i(L_w)]^{\mathbb{T}}$ denotes the measurement time series from Bus i with discrete time index $l_w = 1$ to L_w, and $\mathbb{T}$ denotes matrix transpose. $\boldsymbol{\zeta}(l_w) = [s_1(l_w)\ s_2(l_w)\ \cdots\ s_n(l_w)]^{\mathbb{T}}$ denotes the measurement from Bus 1 to n at discrete time point l_w. The realization of Granger causality estimation is based on vector autoregressive modelling (VAR) of time series. Fitting our test data into a p_cth order VAR(p_c) model gives

$$\boldsymbol{\zeta}(l_w) = \sum_{k_c=1}^{p_c} \mathbf{C}_{k_c} \cdot \boldsymbol{\zeta}(l_w - k_c) + \boldsymbol{\varepsilon}_{l_w}, \tag{3.3.2}$$

where $\mathbf{C}_{k_c}$ are the regression coefficients matrix corresponding to time lag k_c, and $\boldsymbol{\varepsilon}_{l_w}$ are the residuals which are modeled as white Gaussian random vectors. The model order p_c is estimated by the Bayesian information criterion according to the number of samples L_w, number of variables n (number of buses in our case), and number of trials. Numerically, Granger causality evaluates the degree to which the past of the causer variable helps predict the causee variable conditional on the past of the causee itself. Since none of the variables/buses are absolutely independent from other buses in the system under investigation, we cannot simply ignore the other buses while evaluating the causality between a selected pair. Thus to calculate the Granger causality from Bus j to i ($i \neq j$) actually means to calculate the Granger causality from bus j to i conditioned on all the other buses indexed by $\mathbf{r}_{i,j}$, where $\mathbf{r}_{i,j}$ is the vector of all the bus indexes except i and j. In this way, the full regression for the Bus i is recast as

$$s_i(l_w) = \sum_{k_c=1}^{p_c} \Big(\mathbf{C}_{ii,k_c} s_i(l_w - k_c) + \mathbf{C}_{jj,k_c} s_j(l_w - k_c) + \mathbf{C}_{rr,k_c} \mathbf{s}_{\mathbf{r}_{i,j}}(l_w - k_c) \Big) + \boldsymbol{\varepsilon}^i_{l_w}, \quad (3.3.3)$$

where $\mathbf{s}_{\mathbf{r}_{i,j}}[l_w]$ is the measurement vector from the buses indexed by $\mathbf{r}_{i,j}$, $\mathbf{C}_{xx,k_c}$ are the regression coefficient matrices corresponding to time lag k_c for the data component from Bus x. And $\Sigma_{ii} \equiv \mathrm{cov}(\boldsymbol{\varepsilon}^i_{l_w})$ is defined as the covariance of $\boldsymbol{\varepsilon}^i_{l_w}$. To evaluate the impact of Bus j on bus i, we then compute the reduced regression by omitting the $s_j(l_w)$ component from the full regression, which gives

$$s_i(l_w) = \sum_{k_c=1}^{p_c} \Big(\mathbf{C}'_{ii,k_c} s_i(l_w - k_c) + \mathbf{C}'_{rr,k_c} \mathbf{s}_{\mathbf{r}_{i,j}}(l_w - k_c) \Big) + \boldsymbol{\varepsilon}'^i_{l_w}, \quad (3.3.4)$$

and $\Sigma'_{ii} \equiv \mathrm{cov}(\boldsymbol{\varepsilon}'^i_{l_w})$. Then the Granger causality from Bus j to Bus i conditioned on Buses $\mathbf{r}_{i,j}$ is

$$\mathcal{G}_{j \to i|r} \equiv \begin{cases} \ln \frac{|\Sigma'_{ii}|}{|\Sigma_{ii}|} & \text{for } i \neq j \\ 0 & \text{for } i = j \end{cases} \quad (3.3.5)$$

However, the Granger causality focuses merely on the causal relationship between two variables, and the basic property of the signals themselves are ignored, for example the strength of fluctuation of the signal during the fault. Generally, the further a bus locates from the impact source, the weaker the impact this bus might receive, hence the smaller the fluctuation would be. Relying solely on the Granger causality estimation, some buses might be identified as the bus with major impacts due to only local disturbances. To complement to Granger causality the information from strength of signal fluctuations, we propose the fluctuation strength weighted Granger causality. The weighted Granger causality $\mathcal{R}_{ij}$ balances the Granger causality value of each paired variables by the ratio of their variance, which can be expressed as

$$\mathcal{R}_{ij} = \mathcal{G}_{j \to i|r} \cdot \frac{var(\mathbf{s}_j)}{var(\mathbf{s}_i)}. \tag{3.3.6}$$

Then the fault impact causal analysis is accomplished by summarizing the total effect for each bus upon all the other buses, and buses with the largest total effect are identified as the major impact sources.

3.4 Numerical Simulation and Results

In order to demonstrate the big data characterization approach and its associated data-driven application on fault detection and identification, the IEEE 39-bus system (System 1) and IEEE 118-bus system (System 2) are employed for experiments on numerical simulations. All the simulations are executed using a computer with 3.60 GHz Intel i7 CPU and 32 GB RAM. The time consumption for each simulation is less than 10 s. In this chapter, the IEEE 39-bus system and IEEE 118-bus system are modeled in PSCAD to simulate the normal and abnormal operation conditions, which includes the generator grounding, load loss, generator outage, single transmission line outage, and three-phase transmission line outage. Matlab is used to compute the numerical results with the HMM and WGC toolbox [76, 77]

3.4.1 OSMDSA within a Single SVC Zone

In Table 3.2 and 3.3, the OSMDSA results of the proposed approach for System 1 and 2 are demonstrated, respectively. For each SVC zone, the selected optimal synchrophasor measurement devices can be dynamically computed by OSMDSA as shown in Table 3.2 and 3.3. With the dynamic OSMDSA, the full SG can be observed with reduced number of synchrophasor measurement devices. For example, in System 1, the 7 synchrophasor measurement devices located at the pilot buses are selected to determine the SVC zone, if a fault is located in SVC zone III. Then, the synchrophasor measurement devices located at bus 4, 7, 11, 13, 31, and 32 are

Table 3.2: Proposed OSMDSA with SVC zones in System 1

SVC zone	Bus Number	Pilot Bus	Result of Basic OSMDSA	Result of OSMDSA with SLO
I	2, 3, 18, 25, 30, 37	2	2, 3, 25	2, 18, 30, 37
II	1, 9, 39	1	1, 9	1, 9
III	4-8, 10-14, 31, 32	7, 11	4, 7, 11, 13, 31, 32	4, 6, 8, 11, 13, 31, 32
IV	15-17, 19-24, 35, 36	16	16, 20, 22, 33, 36	15-17, 19, 22, 23, 33-36
V	26, 27, 28, 29, 38	28	27, 28, 38	27, 28, 38

Table 3.3: Proposed OSMDSA with SVC zones in System 2

SVC zone	Bus Number	Pilot Bus	Result of Basic OSMDSA	Result of OSMDSA with SLO
I	1-20, 117	2, 13	2, 5, 7, 9, 12, 13, 18, 19	2, 3, 5, 7, 8, 10-12, 15, 17, 19, 20, 117
II	21-23, 25-32, 113-115	114	22, 25, 28, 30, 32, 114	21, 22, 25, 27, 28, 30-32, 113, 114
III	33-48	39, 45	33-35, 37, 41, 45, 46	33, 34, 36-39, 41-43, 45, 47, 48
IV	49-69, 116	57, 67	49, 53, 56, 59, 64, 67, 116	49, 51, 53, 56, 57, 59, 62, 64, 65, 67, 68, 116
V	24, 70-81, 97-99, 118	78,118	24, 71, 74, 78, 80, 118	24, 70, 71, 73, 74, 76, 78, 80, 81, 97-99, 118
VI	82-96, 100-102	88, 93	82, 84, 86, 88, 90, 93, 95, 101	82, 84, 86-88, 90, 92, 94, 95, 101
VII	103-112	108	103, 106, 108, 110	103, 105, 106, 108, 110-112

selected by the basic OSMDSA. Therefore, the total number of selected optimal synchrophasor measurement devices is 11, which is 28.2% of all the synchrophasor measurement devices.

For a simulation period of 12 s, 12000 data samples are collected for system analysis [73]. After 30 MPD iterations for feature extraction, the 12000 data samples are condensed into 30 four-dimensional feature vectors $[\alpha, \tau, \nu, \sigma]^T$, with each representing the amplitude, time-shifting, frequency-shifting and variance of the selected Gaussian atoms. The total data number is 30×4=120, and the data compression rate, which is defined as the data volume generated by the spatial-temporal characterization to the original data volume, is 1.00%. Considering the scenario above, the synchrophasor measurement devices selection rate is 28.2%. Therefore, for this scenario, the total data compression rate can reach 0.282% with controllable residue energy.

3.4.2 OSMDSA within Several SVC Zones

Different from the scenario where faults can be located in a single SVC zone, some faults occur at the adjacent areas between two or several SVC zones. In these scenarios, these SVC zones are combined to form a single SVC zone.

For example, as shown in Fig. 3.3, a transmission line outage occurs between bus 100 and bus 103 between SVC zone VI and SVC zone VII. According to WGC, SVC zone VI and VII are both identified as faulty areas. With the topology information, the optimal synchrophasor measurement devices are selected. Two different synchrophasor measurement devices selection methods are shown in Table 3.4, Method 1 is to select synchrophasor measurement devices in two separate SVC zones, and Method 2 is to select synchrophasor measurement devices in the combined SVC zone, which is the combination of SVC zone VI and VII. Comparing these two methods, the number of selected synchrophasor measurement devices using Method 2 is smaller in both the results of basic OSMDSA and OSMDSA with SLO. It can be expected that the proposed approach can reduce more synchrophasor measurement devices in larger SGs.

Table 3.4: Comparison of Dynamic OSMDSA in Different Scenarios

	Result of OSMDSA	Result of OSMDSA with SLO
	Location	Location
Method 1	82, 84, 86, 88, 90, 93, 95, 101, 103, 106, 108, 110 (12 in total)	82, 84, 86-88, 90, 92, 94, 95, 101, 103, 105, 106, 108, 110-112 (17 in total)
Method 2	82, 84, 86, 88, 90, 92, 95, 100, 105, 110 (10 in total)	82, 84, 86-88, 90, 92, 94, 95, 100, 105, 107, 109, 111, 112 (15 in total)

3.4.3 Fault Detection and Identification

Five representative fault types are employed for evaluating the proposed characterization approach, which are generator grounding, load loss, generator outage, single transmission line outage, and three-phase transmission line outage. The Gaussian atom dictionary is built with 1,800,000 Gaussian atoms and the MPD iteration number is set to 30. For evaluating fault detection and identification performance, 280 and 400 random faults are simulated to generate training and testing data in System 1 and System 2, respectively. An additive white Gaussian noise (AWGN) with SNR 50 dB is added to the measurement for evaluating the performance with noisy measurements [73, 74].

With all the synchrophasor measurement devices selected in System 1 and 2, the results of the fault detection and identification rates are demonstrated in Table 3.5. The result illustrates that the noise impacts heavier in fault identification than detection. Incorporating the OSMDSA results, the fault detection and identification rates are illustrated from Table 3.6.

Compared with Table 3.5, the average detection rates of the proposed approach in Table 3.6 are 100% in the scenarios with noise and without noise, which indicate the proposed approach has high fault detection performance. Considering the scenarios without noise, in Table 3.6, the fault identification rates of the proposed approach are only slightly lower than the fault identification rates in Table 3.5, which indicates the high performance in fault identification. Especially, considering

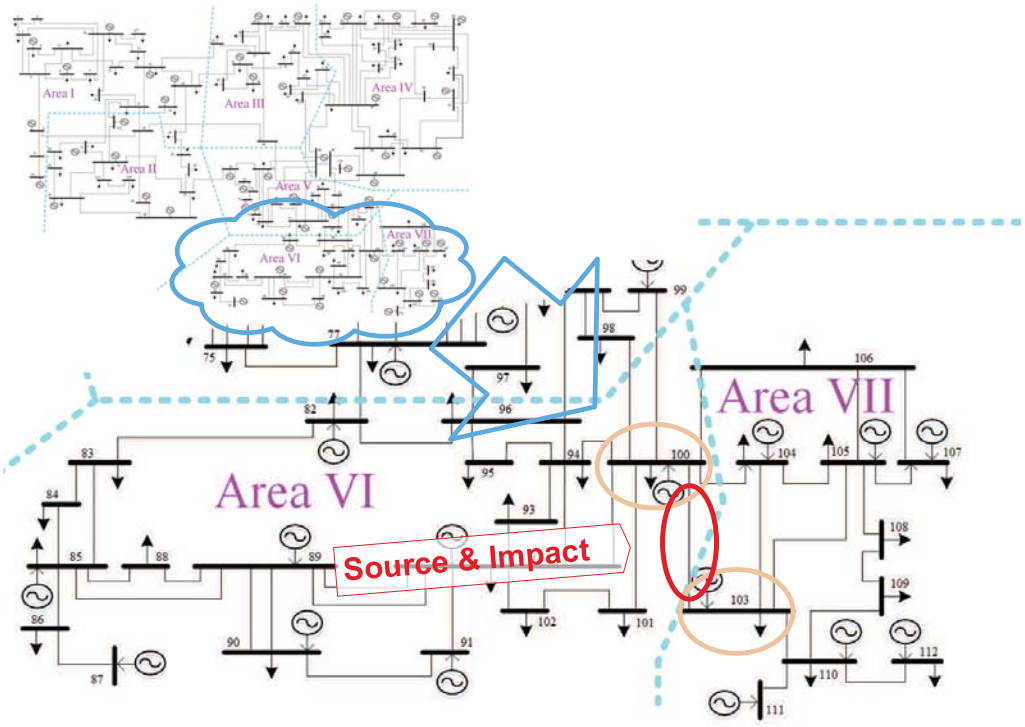


Figure 3.3: An example of fault location in the IEEE 118 bus system.

single line outage, the proposed approach selects more synchrophasor measurement devices, which leads to higher identification rates than the basic OSMDSA. Considering the noise scenarios, in Table 3.6, the identification rates of proposed approach are slightly lower than the identification rates in Table 3.5. However, the maximum difference is smaller than 10%, and the minimum identification rate of the proposed approach is higher than 80%.

3.4.4 Compared with Other Approaches

Data Compression

In [78,79], the optimal synchrophasor measurement devices placement is studied to reduce the number of synchrophasor measurement devices while ensuring the full

Table 3.5: Fault Detection and Identification with Synchrophasor Measurement Devices Fully Placed

Without Noise	Detection	Identification
System 1	100%(280/280)	95.7%(268/280)
System 2	100%(400/400)	96.6%(386/400)
SNR = 50 dB	Detection Rate	Identification Rate
System 1	100%(280/280)	92.8%(260/280)
System 2	100%(400/400)	93.5%(374/400)

Table 3.6: Fault Detection and Identification with Proposed Approach

Without Noise	Detection	Identification	Identification with SLO
System 1	100%(280/280)	90.0%(252/280)	92.1%(258/280)
System 2	100%(400/400)	90.3%(361/400)	92.8%(371/400)
SNR = 50 dB	Detection Rate	Identification Rate	Identification with SLO
System 1	100%(280/280)	82.8%(232/280)	86.8%(243/280)
System 2	100%(400/400)	83.2%(333/400)	87.3%(349/400)

system observability. In this chapter, the dynamic OSMDSA is proposed based on the background that the synchrophasor measurement devices can be widely located in SGs. Combined with SVC zones and pilot buses, the proposed dynamic OSMDSA can provide essential information of the grid with a fewer synchrophasor measurement devices. With MPD, according to the analysis in Section 3.4.1, the proposed approach is able to decrease the data compression rate to 0.282% according to the results in Table 3.2, 3.3 and 3.4. Compared with the optimal synchrophasor measurement devices placement, the proposed OSMDSA is more effective and flexible in big data processing and feature extraction.

Fault Detection and Identification

In [22, 23], ANN and SVM are studied to detect and diagnose faults in SGs. ANN is a supervised learning model with hidden neurons in different layers, which can be used for fault classification and pattern recognition. SVM is also a supervised learning model, which builds a hyperplane with different kernels for feature classification. As shown in Table 3.7, compared with the result with ANN and SVM, the proposed approach has the best performance in detection and identification. As shown in Table 3.8, the proposed approach has the shortest time consumption because it provides a significant data compression rate, which dramatically reduces the computation load.

Fault Location

Compared with the model based sensitivity analysis approach in [80–82], the proposed WGC analysis is an effective and accurate approach for fault location. The model based sensitivity analysis requires the detailed dynamic model of the SGs, topology information of the SGs, and it also incurs large computation load and long simulation time. Although SG can be accurately modeled, however, in most real-world situations large-scale SGs can not be accurately modeled due to lack and change of system information and parameters, especially under time-varying operating conditions. Our proposed method can be used for analysis of causality without accurate information of SG typology and system parameters. And thus, it is valuable in security assessment with partial system information. Combined with the MPD, the proposed WGC can locate the fault precisely under noise-effected scenarios. In our simulations, 60 faults are simulated at random locations in the two test systems with noise or without noise, the successful fault location rate is 100%.

Table 3.7: Performance Comparison of Different Methods with Reduced Number of Synchrophasor Measurement Devices

	Detection	Identification	Identification, 50 dB
ANN	94.2%	83.6%	78.7%
SVM	95.6%	81.3%	75.5%
Proposed Method	100.0%	92.4%	87.1%

Table 3.8: Time Consumption Comparison of Different Methods with Reduced Number of Synchrophasor Measurement Devices

	System 1 (S)	System 2 (S)
ANN	470.2	1482.6
SVM	590.2	1924.7
Proposed Method	2.7	3.6

Further, even for large-scale SG which can be accurately modeled, the proposed WGC method provides an approach for rapid and rough determination of event cause, which will provide valuable apriori information for subsequent security assessment and contingency analysis. The above reasons justify our novel application of the WGC approach.

3.5 Chapter Conclusion

To respond to the big data challenge in modern SGs, we aim to provide our initial considerations, investigation, and to propose a fault analysis solution based on characterization and analytics of the synchrophasor measurements.

The main contributions of this chapter can be concluded as: the proposed two-layer dynamic OSMDSA provides the spatial characterization of the synchropahsor measurement system, and MPD is used for characterizing the signals in the temporal-domain. Combing the spatial-domain and temporal-domain characteri-

zations, the volume of the Big Data is substantially reduced, while preserving the comprehensive information in the data. To evaluate the proposed approach, a HMM based SG situational awareness method is proposed to investigate the fault detection and identification; the WGC for SGs is proposed to analyze the causal relationship of the buses during system disturbance.

The numerical results and comparison with other methods demonstrate that, although the data volume is reduced below 1% of the original data volume, the situational awareness performance was only slightly affected, and this fact demonstrate the effectiveness and robustness of the proposed analytics approach in synchrophasor based *Big Data*.

Chapter 4

Social Energy Mining in Distribution Grids

The inherent nature of energy, physicality, sociality and informality, determined that development of power systems would eventually come to a point, that active informative interactions between the society and the power system will be an inevitable part for further improvements. And today's cheap sensor manufacturing, maturing big data analyses and boost of intelligent techniques have paved the way to this revolutionary point to transfer the existing power system to a social energy system.

A social energy system is a complex of physical energy systems, physical social systems, and the artificial virtual systems derived from the physical systems. The artificial virtual systems are derived with certain purposes that concern the joint operation of the socio-technical systems. Utilizing the multifaceted data collected from the socio-technical systems, through sufficient interacting and massive computing, knowledge automation of the systems is gained, and intelligence in system control and management is generated. The knowledge and intelligence in turn are applied in the social energy system, achieving a truely automated and intelligent joint socio-technical system design and management.

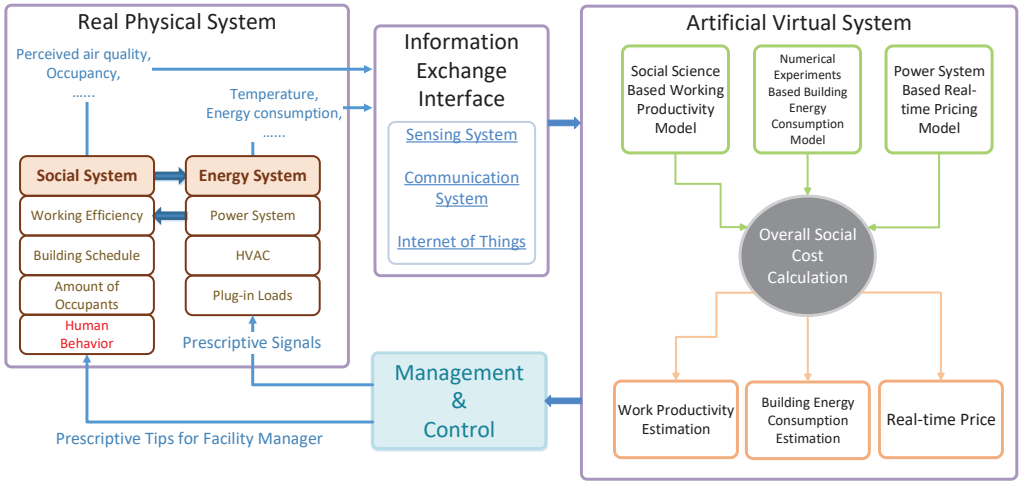


Figure 4.1: Technical scheme of the DU social energy system case study.

In this chapter, we provide a case study on a campus power grid to demonstrate the concept of social energy. The case study includes the elements of power system operation, smart building modeling, real-time pricing mechanism, and human behavior modeling. The technical scheme of the case study is demonstrated in Fig. 4.1, using the concept of parallel intelligence and control. The detailed technical description of the case study is introduced in this section.

In previous literature, a lot of studies have proved the relationship between comfort and work performance. Indoor temperature is a critical factor of the indoor environment, which can affect human behavior in many ways such as perceived air quality, working performance, thermal comfort, etc. [56] The U.S. federal government regulates CO_2 emission for universities, as a result, university facility managers are required to meet the green house reduction regulation in order to avoid financial penalties by targeting on energy cutback. However, some of the strategies are not efficient and may cause reduction in comfort. The occupants who feel uncomfortable are less productive need more time to accomplish their tasks, which may lead to more energy consumption and environmental degradation [83]. [57, 84] show that with $2 saving per employer when indoor temperature is within comfortable range,

and the working efficiency will reduce 2% per degree Celsius when the temperature is higher than 25°C.

This case study proposes a smart building power consumption strategy by jointly considering the interactions between the campus power grid and the community artificial system. Six buildings in the University of Denver are chosen as the target buildings for demonstration, and Table 4.1 provides the basic information about them. Work productivity is considered as one of the essential factors in the methodology as work performance varies considerably under different indoor temperatures.

Table 4.1: Basic Information about the 6 target buildings

	Name	Building Type	Area (ft^2)	Floor
1	Ritchie Center	Fitness center	440000	4
2	Law Building	Academic	181000	4
3	Sturm Hall	Academic	245000	4
4	Daniels Building	Academic	110536	6
5	Newman Center	Performing art	180000	5
6	Olin Hall	Academic	40000	2

The case study is divided into two parts regarding the power consumption model of the target buildings. We started the case study under a simpler condition, where the consumption of each hour of the day is independent from each other, more specifically the buildings' power consumption is solely a function of the indoor temperature settings and outdoor temperatures of the current time $e_t = h(T_t^{in}, T_t^{out}, t)$. And indoor temperature T_t^{in} is the management variable, and the hourly lowest overall costs of the system is set as the management object. We expanded the scenario into a more realistic and complex one in the second part of the case study by replacing the static consumption model with a dynamic consumption model. In this scenario, consumption, outdoor temperature and hour together form the state of the system at time t, which is $x_t = [e_t, T_t^{out}, t]^\intercal$. And consumption e_t is not only a function of outdoor temperature T_t^{out} and the control input indoor temperature

setting T_t^{in}, but also a function of previous consumption e_{t-1}. And the minimization of 24-hour aggregated overall costs is set as the management object. Neural Networks are used in both parts to learn and train the consumption profile models for each building. And training data are generated by the smart building models, which are built in the software eQUEST [85].

It should be noted that distribution locational marginal real-time pricing is implemented to reflect the changing of energy usage and utility expenses. The DLMP calculation is developed in MATLAB environment based on MATPOWER simulation toolbox [86]. Given the human work performance model, energy consumption profile models and the DLMP, distributed iterative adaptive programming [87] is applied to find out the optimal indoor temperature settings for the 6 target buildings in the first part, and Neural Network implementation of iterative adaptive dynamic programming [88] is introduced to solve the optimization problem.

4.1 Case study of social energy implementation in a campus grid with static building energy consumption model

The technical artificial part of the campus social energy system described in Fig. 4.1 is composed of three key models. This section will give detailed description of the three models, and how they cooperate and interact with each other to guide the energy management system.

4.1.1 Human Work Performance Model

Indoor temperature is one of the fundamental characteristics of the indoor environment. It can be controlled with a degree of accuracy depending on the building and its HVAC system. The indoor temperature affects thermal comfort, perceived

air quality, sick building syndrome symptoms and performance at work [57]. In this study, the work productivity P is referred to the effects of temperature on performance at office work [56], and the human work performance model can be expressed as

$$\begin{aligned} \xi &= g(T^{in}) \\ &= 0.1647524 \cdot (\frac{5}{9} \cdot (T^{in} - 32)) - 0.0058274 \cdot (\frac{5}{9} \cdot (T^{in} - 32))^2 + \\ &\quad 0.0000623 \cdot (\frac{5}{9} \cdot (T^{in} - 32))^3 - 0.4685328 \end{aligned} \tag{4.1.1}$$

where ξ is the work productivity, T^{in} is the indoor temperature settings which satisfies $64 \leqslant T^{in} \leqslant 79$. And Fig. 4.2 shows the relationship between indoor temperatures and the corresponding work efficiency. It should be pointed out that, although according to [89], the ideal temperature range for university buildings is between 68°F and 74°F, the temperature bracket in our study is designed from 64°F and 79°F, as the target buildings include a fitness center.

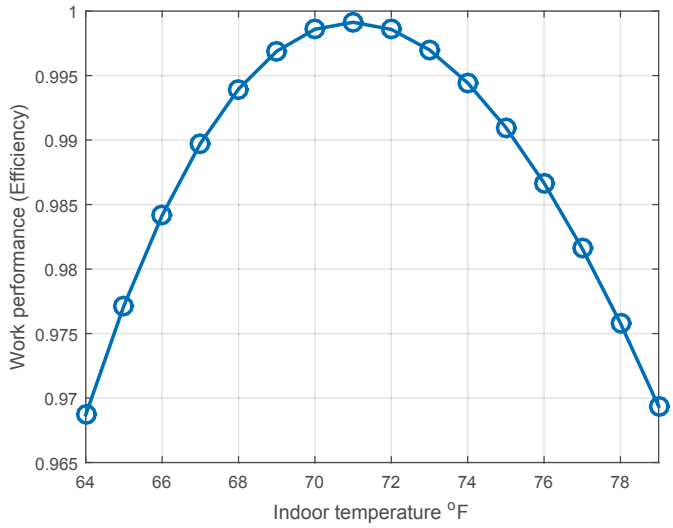


Figure 4.2: Work efficiency as a function of indoor temperature.

Let $\xi_{k,t}$ and $u_{k,t}$ denotes the work productivity and indoor temperature setting in building k at time t, respectively, (4.1.1) can be rewritten as

$$\xi_{k,t} = g(u_{k,t}). \tag{4.1.2}$$

In addition, we denote $\boldsymbol{\xi}_t = [\xi_{1,t}\ \xi_{2,t} \cdots \xi_{6,t}]^\intercal$ and $\mathbf{u}_t = [u_{1,t}\ u_{2,t} \cdots u_{6,t}]^\intercal$, $\mathbf{x}_t$ is the control variable in this case study.

4.1.2 The Neural Network Based Energy Consumption Profile Models

We utilize eQUEST [85] as the building simulation tool, which can provide comprehensive and detailed calculations about HVAC systems and simplistic assumptions for lighting and plug loads. The hourly report from the simulation system can provide sufficient information for training the neural network models for predicting power consumption. Table 4.2 shows partial simulation results of the hourly energy consumption on July 1*st* for the Ritchie center. It is noted that the simulation generated much more information and results, however, due to space limitation only the following is included in this chapter.

Table 4.2: Partial simulation results of Ritchie center on July 1*st* from 15 : 00 to 20 : 00

Hour	T_{in} (°F)	T_{out} (°F)	Energy (BTU)
15	64	94	1.08×10^7
16	64	92	1.31×10^7
17	64	92	1.70×10^7
18	64	93	2.03×10^7
19	64	89	2.22×10^7
20	64	85	2.19×10^7

Simulated data of one entire year with constant indoor temperature setting at 64, 67, 70, 73, 76 and 79 °F were generated for all the 6 buildings, and the sample

data on all workdays were picked for training the neural network energy consumption models. For each individual building, a two-layer feed-forward network with sigmoid hidden neurons and linear output neurons is trained with the Levenberg-Marquardt backpropagation algorithm. Hour, indoor temperature setting and outdoor temperature are the inputs, and energy consumption is the output of the neural network model. 75% of the sample data are randomly selected as training data, 15% as validation data and 15% as testing data. Table 4.3 lists the number of hidden neurons for training the consumption model of each building and the regression R-value of the training results. The average R-value is 0.92, which demonstrates that the 6 energy consumption neural network models are acceptable. Thus, energy consumptions of building k at time t, which is denoted as $e_{k,t}$ can be expressed as functions of indoor temperature setting $u_{k,t}$, time t and outdoor temperature T_t^{out}

$$e_{k,t} = h_k(u_{k,t}, t, T_t^{out}) \tag{4.1.3}$$
$$\mathbf{e}_t = H(\mathbf{u}_t, t, T_t^{out})$$

where $\mathbf{e}_t = [e_{1,t}\ e_{2,t} \cdots e_{6,t}]^\intercal$.

Table 4.3: Number of hidden neurons for training the target buildings and the regression R-value of the training results

Index	Building	Hidden Layer	R-value
1	Ritchie	30	0.88
2	Law	20	0.96
3	Sturm	30	0.96
4	Daniels	30	0.94
5	Newman	50	0.86
6	Olin	5	0.94

4.1.3 Distribution Locational Marginal Pricing for DU Campus Grid

To indicate the influence of load changing on both financial aspect and campus power system side, distribution locational marginal pricing is introduced to associate the physical system to the artificial system. The detailed campus power system configurations is depicted in [90]. Fig. 4.3 shows the topology of DU campus power system. Using a power system simulator developed based on DU campus power system, the AC-OPF based DLMP is introduced to implement the real-time pricing mechanism.

$$\arg\max_{p_j^b, p_i^g} \quad z = \sum_{j=1}^{N} (c_j - p_j^b) \cdot q_{c_j} \tag{4.1.4}$$

$$- \sum_{i=1}^{M} (p_i^g - s_i) \cdot q_{s_i}$$

$$\text{s.t.} \quad \sum_{i=1}^{M} q_{s_i} - \sum_{j=1}^{N} q_{c_j} - L_P(V, \theta) = 0 \tag{4.1.5}$$

$$\sum_{i=1}^{M} Q_{s_i} - \sum_{j=1}^{N} Q_{c_j} - L_Q(V, \theta) = 0 \tag{4.1.6}$$

$$f_j(V, \theta) \leqslant f_j^{Max} \tag{4.1.7}$$

$$q_{s_i}^{MIN} \leqslant q_{s_i} \leqslant q_{s_i}^{MAX} \tag{4.1.8}$$

$$Q_{s_i}^{MIN} \leqslant Q_{s_i} \leqslant Q_{s_i}^{MAX} \tag{4.1.9}$$

$$V_i^{MIN} \leqslant V_i \leqslant V_i^{MAX} \tag{4.1.10}$$

where z is the system social surplus that is gained from our DLMP calculation, $N = 57$ is the total number of campus buildings and j is the index of buildings; $M = 3$ is the total number of electricity suppliers and i is the index of those generators; c_j stands for the building bid price for each power generation and s_i represents the offer price from each power generation; p_j^b is the distribution locational

marginal price at each building j, and p_i^g stands for the distribution locational marginal price at supply bus i; q_{c_j} is the power demand at building j; q_{s_i} is the power supply from bus i; V and θ are voltage magnitude and angle, respectively; f_j stands for the power flow at jth line, which is limited by $f_j^{max} = 400$ A; q_{s_i} is the active power output from each power source and the maximum capacity $q_{s_i}^{MAX} = 12.7$ MW, while Q_{s_i} is the reactive power output from the corresponding energy generation and the maximum capacity $Q_{s_i}^{MAX} = 11$ MVar; V_i stands for the voltage magnitude of the ith bus with power injection, in this case study $V_i^{MIN} = 0.90$ pu and $V_i^{MAX} = 1.10$ pu; and $L_P(V, \theta)$ and $L_Q(V, \theta)$ are the total active power loss and reactive power loss in the DU campus power gird, respectively.

For the convenience of later calculation, we express the DLMP of the target buildings $\mathbf{p}_t = [p_{1,t}\ p_{2,t} \cdots p_{6,t}]^\intercal$ simply as a highly nonlinear and complex function $\Gamma(\cdot)$ of energy consumption $\mathbf{e}_t$, and thus as a function of the control variable $\mathbf{u}_t$ as

$$\mathbf{p}_t = \Gamma(\mathbf{e}_t) = \Gamma(H(\mathbf{u}_t, t, T_t^{out})). \tag{4.1.11}$$

4.1.4 Overall Social Cost

To address the overall social cost, a novel method is demonstrated in this section. Based on the aforementioned system configurations and social energy methodology, the formulation for the overall social cost comprises two major parts: the utility cost, which is calculated by the end-use energy and the corresponding DLMPs; the cost of work productivity, which is determined by the cost of performance reduction and the amount of working personnels. (4.1.12) defines the formulation for calculating the overall social costs ψ_t at time t of the target buildings in University of Denver's campus grid.

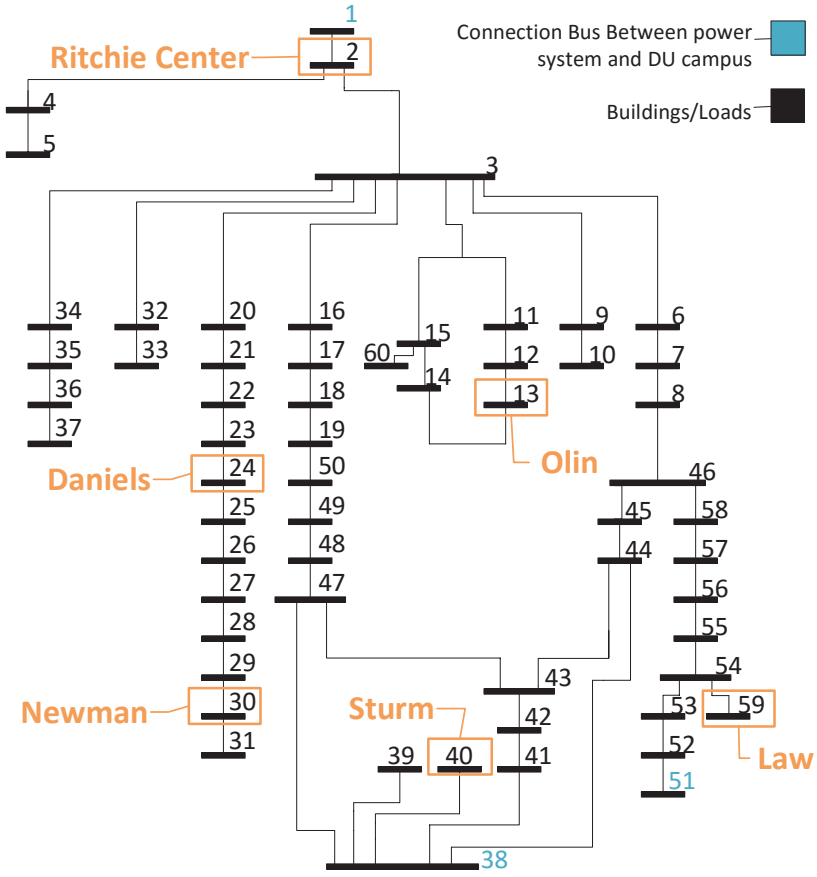


Figure 4.3: The network topology of DU campus grid.

$$\psi_t = \sum_{k=1}^{6}[p_{k,t} \cdot e_{k,t} + \alpha(1 - \xi_{k,t}) \cdot o_{k,t}] \tag{4.1.12}$$

$$= \mathbf{p}_t \cdot \mathbf{e}_t + \alpha(\mathbf{1} - \boldsymbol{\xi}_t) \cdot \mathbf{o}_t \tag{4.1.13}$$

$$= \Gamma[H(\mathbf{u}_t, t, T_t^{out})] \cdot H(\mathbf{u}_t, t, T_t^{out}) + \alpha[1 - g(\mathbf{u}_t)] \cdot \mathbf{o}_t$$

$$= \Psi(\mathbf{u}_t, t, T_t^{out}, \mathbf{o}_t) \tag{4.1.14}$$

where ψ_t is the overall cost at time t, α is the annual saving for each personnel when the working productivity is 1 it is in the unit of dollars, and $\mathbf{o}_t = [o_{1,t}\ o_{2,t} \cdots o_{6,t}]^\intercal$ where $o_{k,t}$ is the number of occupants in building k at time t. It should be noted that,

the outdoor temperature T_t^{out} and number of occupants $\mathbf{o}_t$ are obtained through learning historical data and are directly related to the time t in a day. For this reason, at a certain time point t, parameters t, T_t^{out} and $\mathbf{o}_t$ are known and they are no more variables. Hence the overall social cost is solely the function of the indoor temperature settings $\mathbf{u}_t$. The goal of the case study is to find, at time point t, the best indoor temperature settings $\hat{\mathbf{u}}_t$ which can generate the most economic combination between HVAC costs and work productivity. Therefore, the problem can be formulated as the following

$$\begin{aligned} \hat{\mathbf{u}}_t &= \arg\min_{\mathbf{u}_t} \Psi(\mathbf{u}_t) \\ &\text{s.t.} \quad 64 \leqslant u_{k,t} \leqslant 79 \end{aligned} \tag{4.1.15}$$

4.1.5 Distributed iterative adaptive programing for solving the minimum social cost

Searching in the vector space of $\mathbf{u}_t = [u_{1,t}\, u_{2,t} \cdots u_{k,t}]^\intercal$ can be inefficient and computationally costly, especially when the number of target buildings increases. Therefore, we chose the distributed iterative adaptive programing (DIAP) [87] to solve (4.1.15). Simply speaking, the DIAP algorithm updates the control variable of only one building at a time, and iterates through all the buildings until the control sequence stops changing significantly. Details of the procedures are narrated as the following.

Define $\mathcal{K} = \{1, 2, \cdots, 6\}$ to be the set of building indices. Denote $\mathbf{u}_{\bar{\eta},t} = \{u_{k,t} | k \in \mathcal{K}, k \neq \eta\}$, then the object cost function can be written as $\Psi(u_{\eta,t}, \mathbf{u}_{\bar{\eta},t})$. Let $l = 0, 1, 2, \cdots$ be the iteration index, and $\{\eta_l\}$ be a sequence which satisfies $\eta_l \in \mathcal{K}$, for $\forall l = 0, 1, 2, \cdots$. At the beginning of iteration when $l = 0$, construct the initial control sequence $\mathbf{u}_{\eta_0,t} = [u_{1,t}^0\, u_{2,t}^0 \cdots u_{6,t}^0]^\intercal$, and for $\forall t = 1, 2, \cdots, 24$ the iteration starts as

$$\tilde{u}_{\eta_0,t} = \arg\min_{u^0_{\eta_0},t} \Psi(u^0_{\eta_0,t}, \mathbf{u}^0_{\bar{\eta}_0,t}), \tag{4.1.16}$$

$$\tilde{\psi}_{\eta_0,t} = \Psi(\tilde{u}_{\eta_0,t}, \mathbf{u}^0_{\bar{\eta}_0,t}). \tag{4.1.17}$$

Construct a new control sequence based on (4.1.16) and (4.1.17) as $\mathbf{u}_{\eta_1,t} = [u^1_{1,t}\, u^1_{2,t} \cdots u^1_{6,t}]^\intercal$, where $u^1_{k,t} = u^0_{k,t}$ for $k \neq \eta_0$, and $u^1_{k,t} = \tilde{u}_{\eta_0,t}$ for $k = \eta_0$. Then for $\forall l = 1, 2, 3, \cdots$ and the corresponding η_l, the DIAP algorithm will continue as

$$\tilde{u}_{\eta_l,t} = \arg\min_{u^l_{\eta_l},t} \Psi(u^l_{\eta_l,t}, \mathbf{u}^l_{\bar{\eta}_l,t}), \tag{4.1.18}$$

$$\tilde{\psi}_{\eta_l,t} = \Psi(\tilde{u}_{\eta_l,t}, \mathbf{u}^l_{\bar{\eta}_l,t}). \tag{4.1.19}$$

And construct the next new control sequence $\mathbf{u}_{\eta_{l+1},t} = [u^{l+1}_{1,t}\, u^{l+1}_{2,t} \cdots u^{l+1}_{6,t}]^\intercal$, where $u^{l+1}_{k,t} = u^l_{k,t}$ for $k \neq \eta_l$, and $u^{l+1}_{k,t} = \tilde{u}_{\eta_l,t}$ for $k = \eta_l$.

The iteration stops at $l = L$, when $|\tilde{\psi}_{\eta_L,t} - \tilde{\psi}_{\eta_{L-\delta},t}| \leqslant \beta_1$ and $|\mathbf{u}_{\eta_L,t} - \mathbf{u}_{\eta_{L-\delta},t}| \leqslant \beta_2$, where $\delta = \{1, 2, \cdots, 5\}$, β_1, β_2 are threshold constraints. Thus the optimal indoor temperature settings at time t is $\hat{\mathbf{u}}_t = \mathbf{u}_{\eta_L,t}$, and the minimum social cost is $\hat{\psi}_t = \tilde{\psi}_{\eta_L,t}$. Detailed numerical results are provided in the following section.

4.1.6 Results and Discussion

We selected one typical summer week, July 25th to 29th, and one typical winter week, Feb 22nd to 26th, as the test period to demonstrate the concept of social energy. Figs.4.4 to 4.9 show the results in the summer week, and Figs.4.11 to 4.16 are results for the winter week.

In each of these figures, the plot at the down-right corner displays the outdoor temperature changes in each day (left axis), and the number of occupants with respect to time (thick light blue curve to the right axis). Here we assume that the

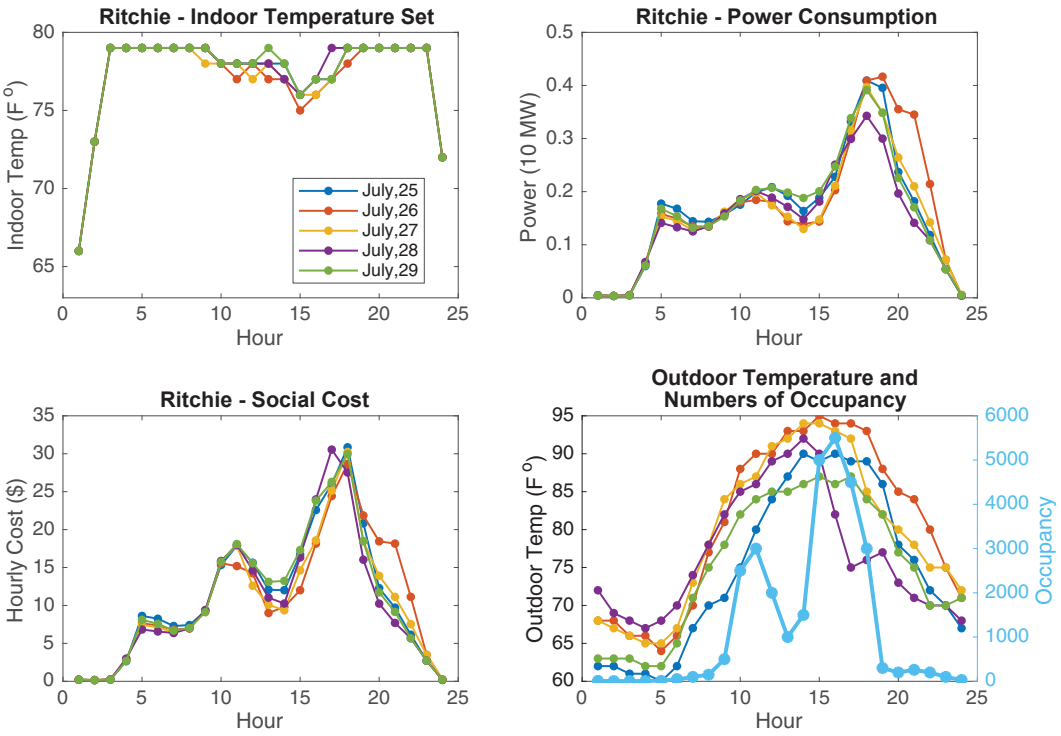


Figure 4.4: Test conditions and results for Ritchie center in a typical summer week.

occupancy change is the same everyday, in each building, respectively. The up-left plot shows the optimized indoor temperature settings calculated by the DIAP algorithm, and the up-right and down-left plots demonstrate the corresponding power consumptions and hourly costs corresponding to the optimized temperature settings, respectively.

We can observe from these figures that the change of hourly cost does not follow exactly the pattern of the power consumption change, and the cost is affected by both the power consumption and the occupancy number's change. For example, in Fig. 4.6 and 4.7, at $13:00$ there is a clear drop in cost caused by the dramatic drop of occupancy. And in Fig. 4.8, the cost curves receive much stronger affects from the occupancy change than from the consumption change.

The overall social cost of the 6 target buildings during the test weeks are provided in Fig. 4.10 and Fig. 4.17. In these two figures, solid lines with "•" markers are the

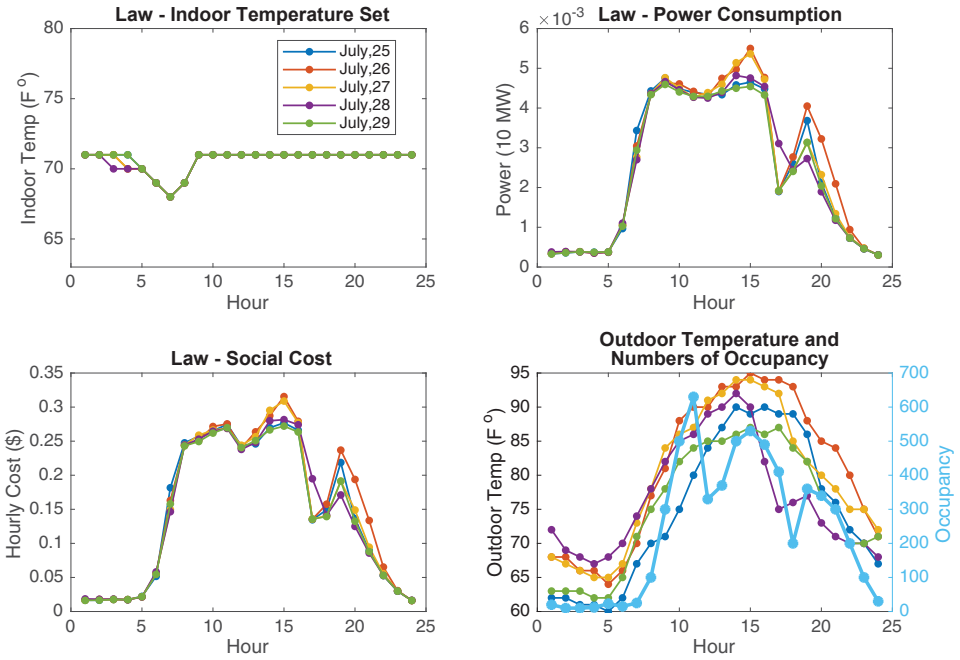


Figure 4.5: Test conditions and results for the Law building in a typical summer week.

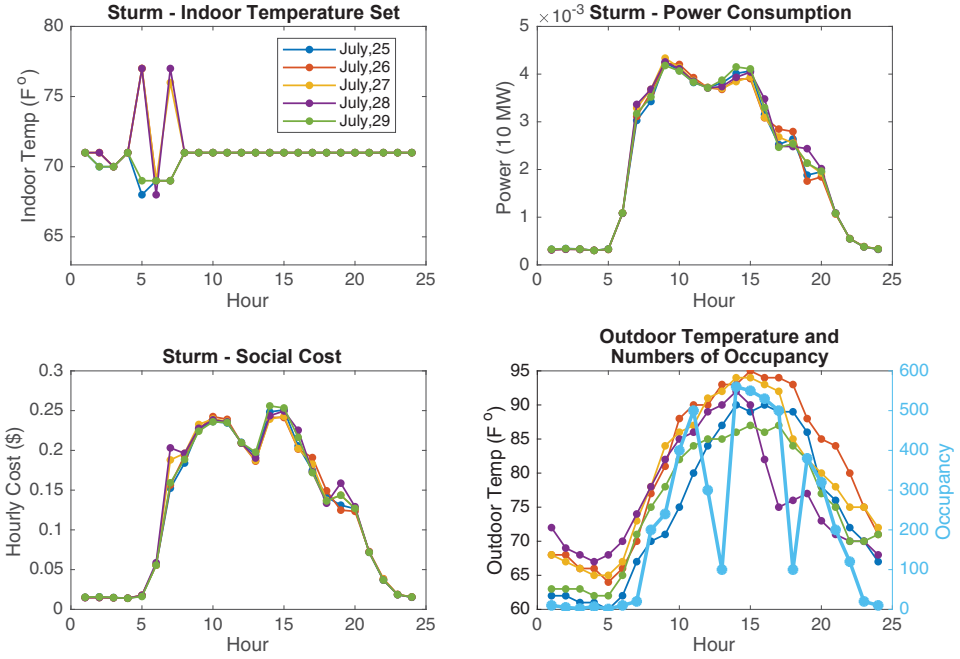


Figure 4.6: Test conditions and results for Sturm Hall in a typical summer week.

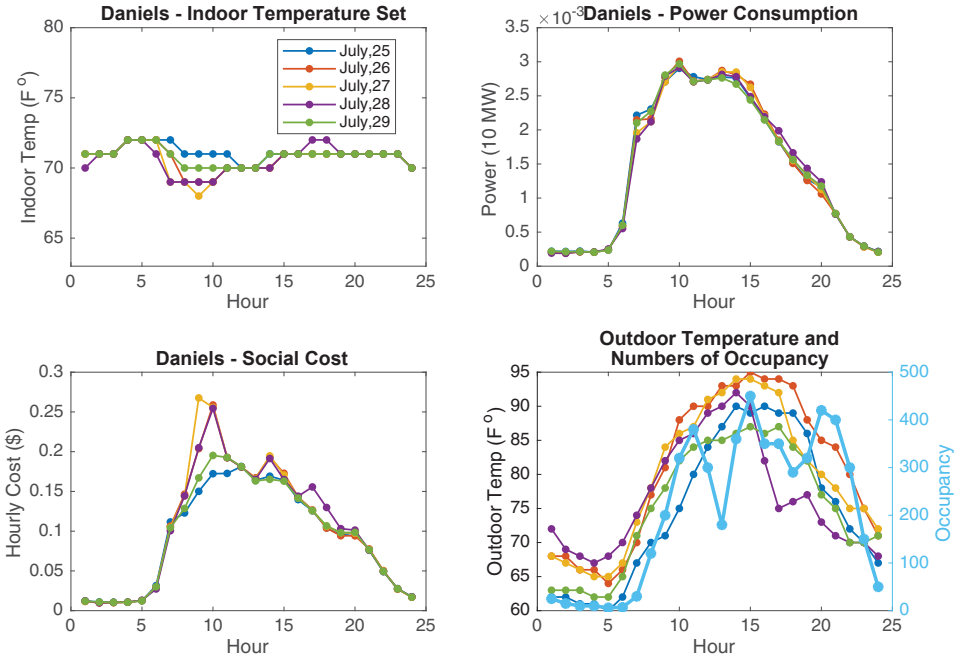


Figure 4.7: Test conditions and results for the Daniels building in a typical summer week.

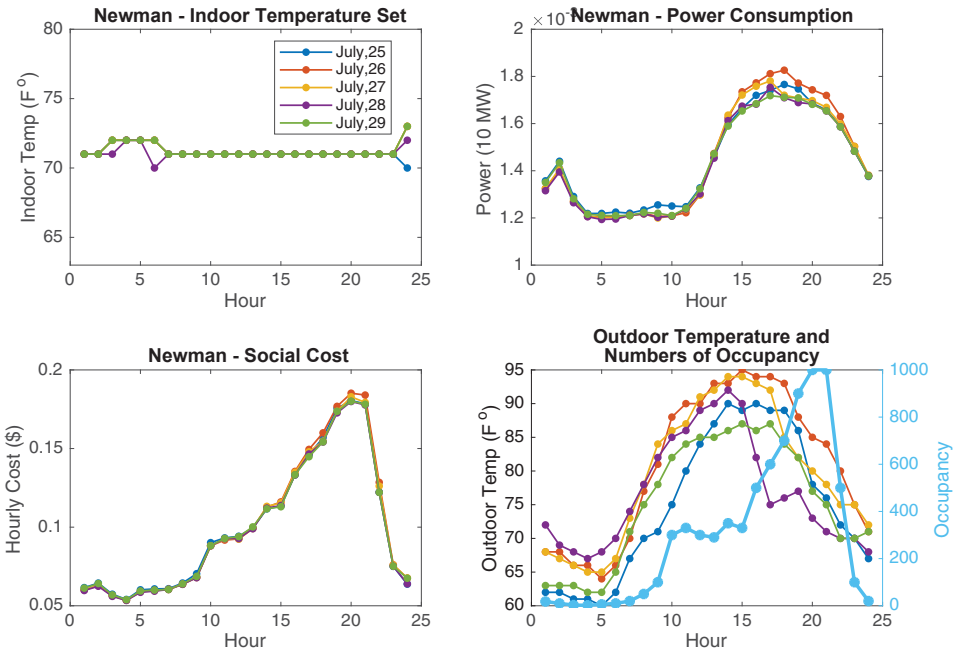


Figure 4.8: Test conditions and results for Newman center in a typical summer week.

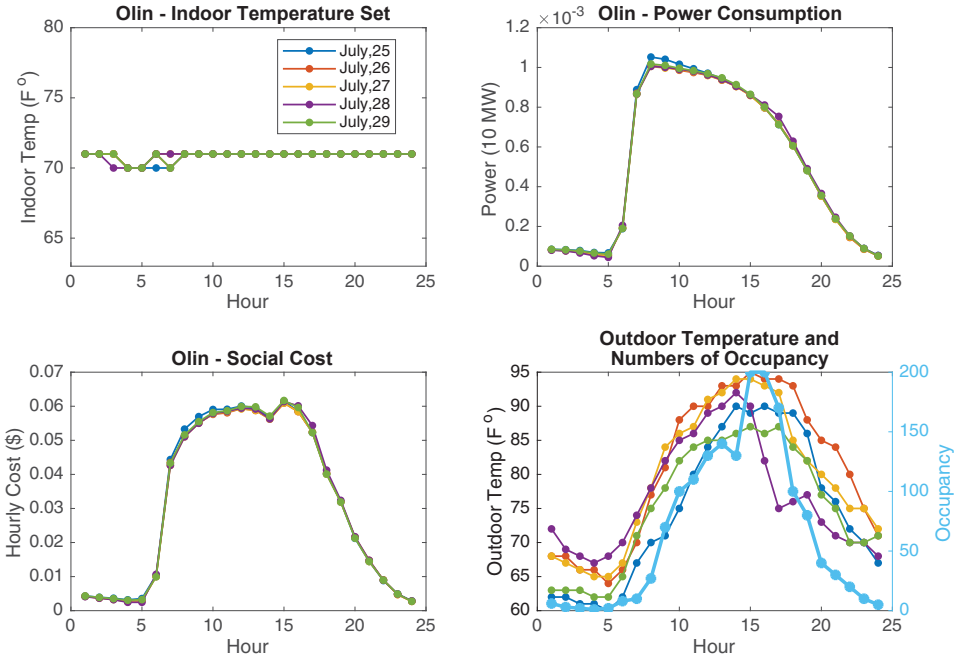


Figure 4.9: Test conditions and results for Olin Hall in a typical summer week.

total hourly costs of the target buildings applying the optimized indoor temperature settings, and different colors denote different days. What's more, we also provide some more test results with 3 baseline reference temperature settings for comparison. The dotted lines with markers "+", "×" and "○" display the total costs when the indoor temperatures are set to 64°F, 71°F and 79°F, respectively. Comparing between the curves with same color but different markers, it is easy to find that the solid lines are always the lowest.

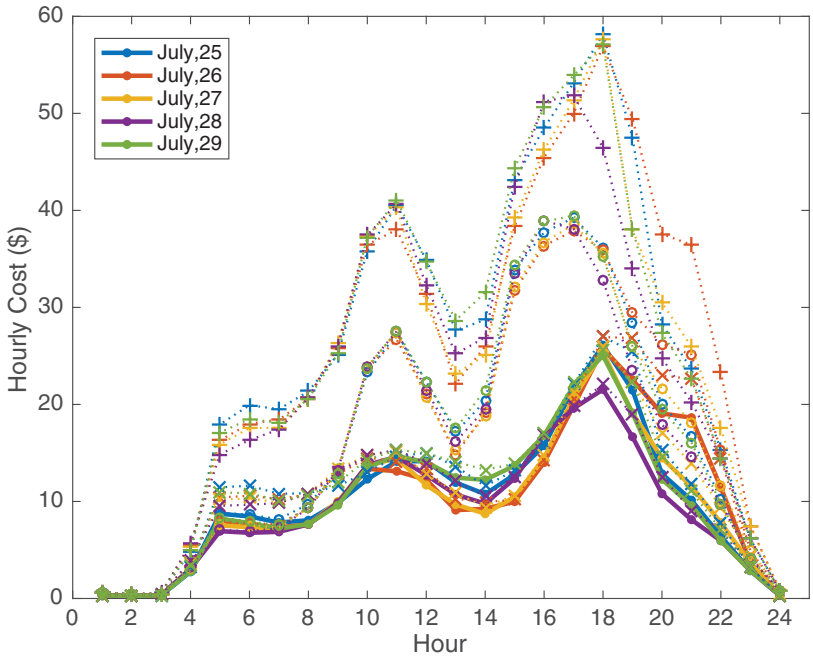


Figure 4.10: Comparison of the overall social cost of the 6 target buildings in the test summer week with different indoor temperature settings.

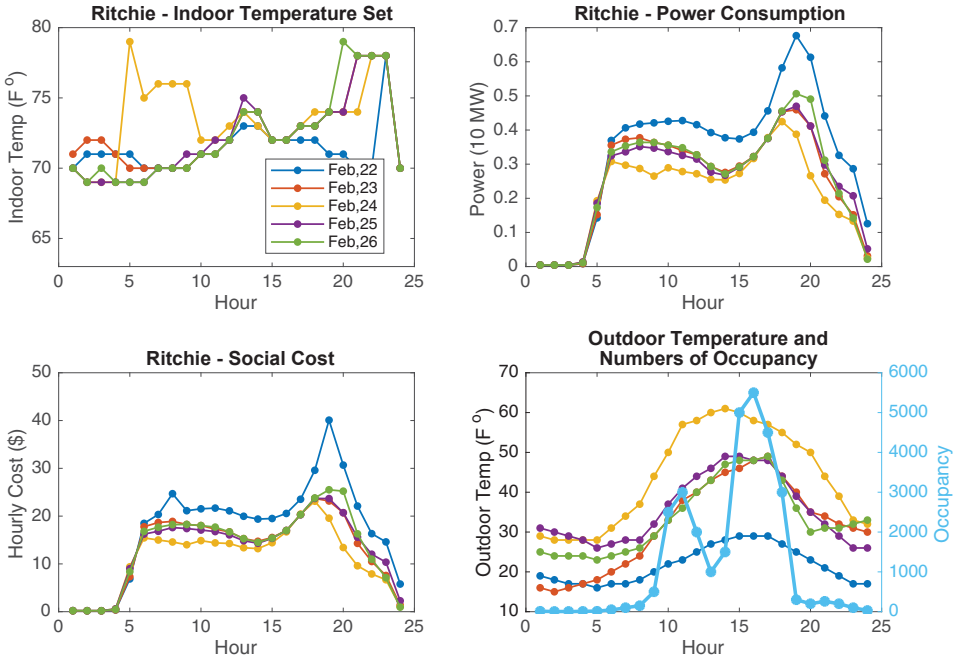


Figure 4.11: Test conditions and results for Ritchie center in a typical winter week.

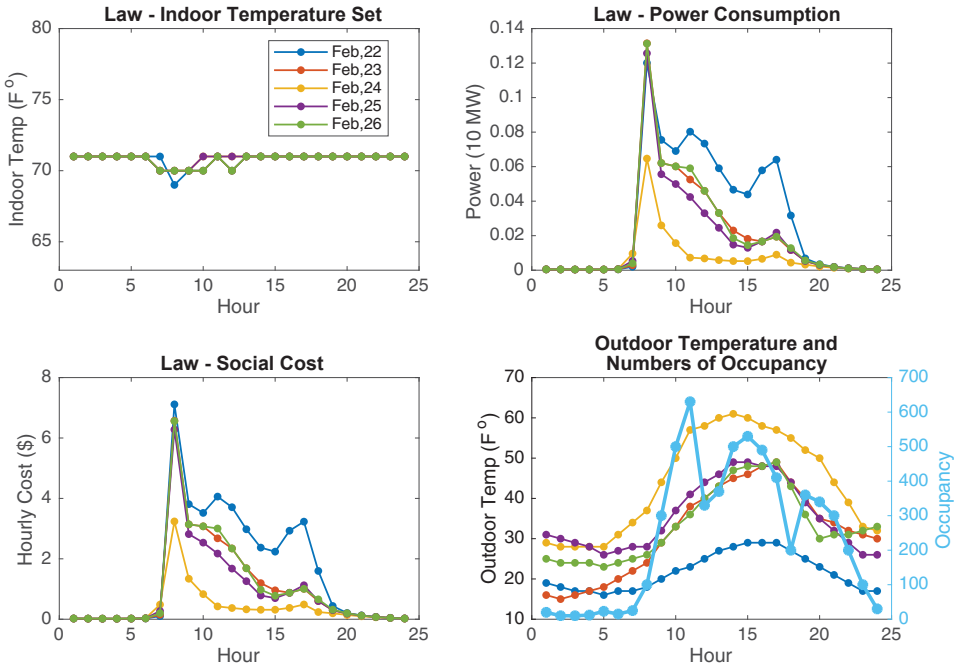


Figure 4.12: Test conditions and results for the Law building in a typical winter week.

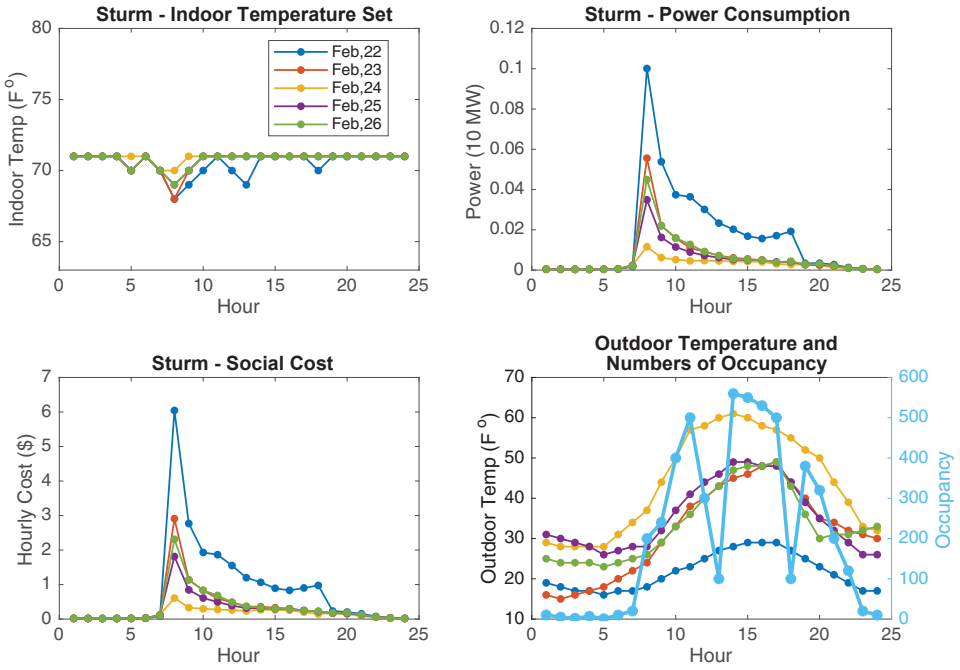


Figure 4.13: Test conditions and results for Sturm Hall in a typical winter week.

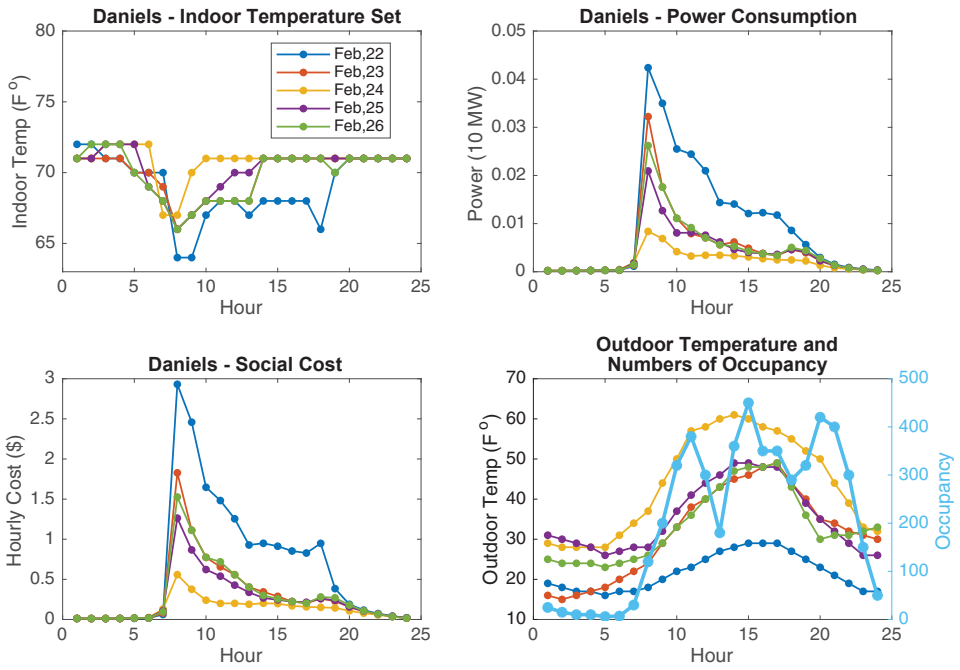


Figure 4.14: Test conditions and results for the Daniels building in a typical winter week.

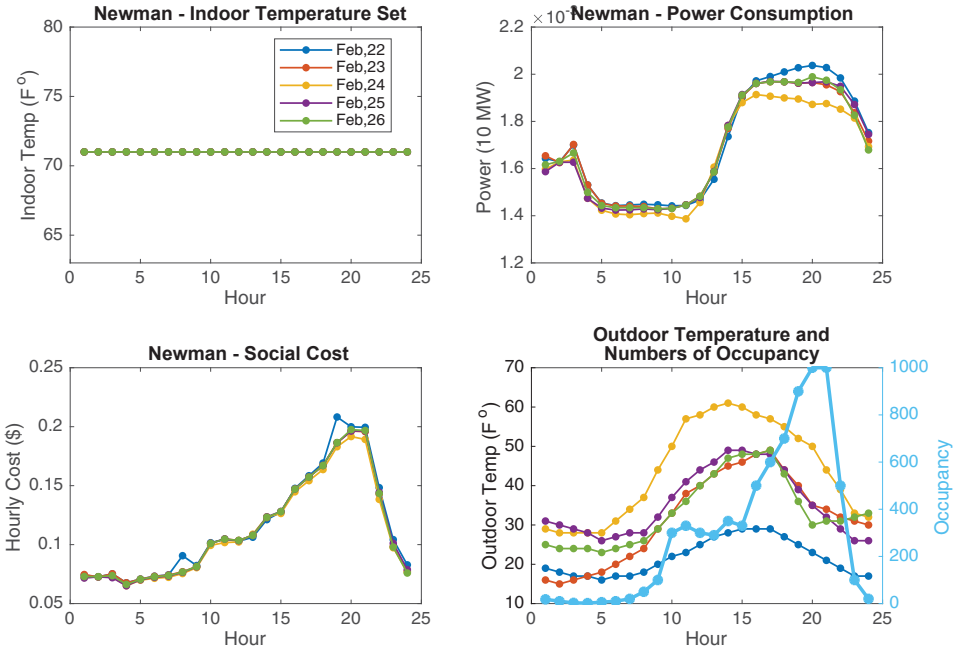


Figure 4.15: Test conditions and results for Newman center in a typical winter week.

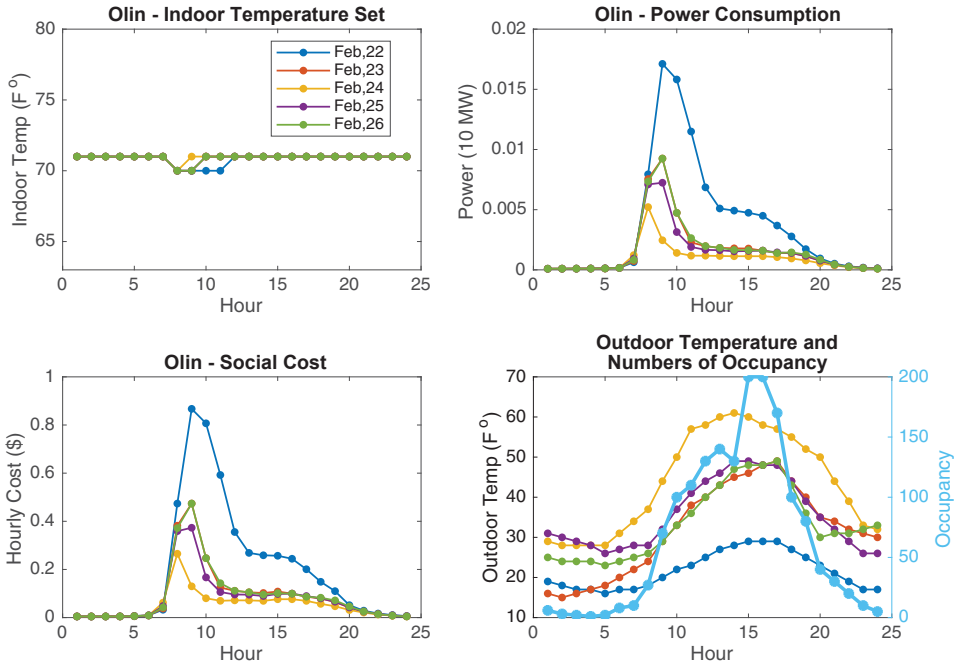


Figure 4.16: Test conditions and results for Olin Hall in a typical winter week.

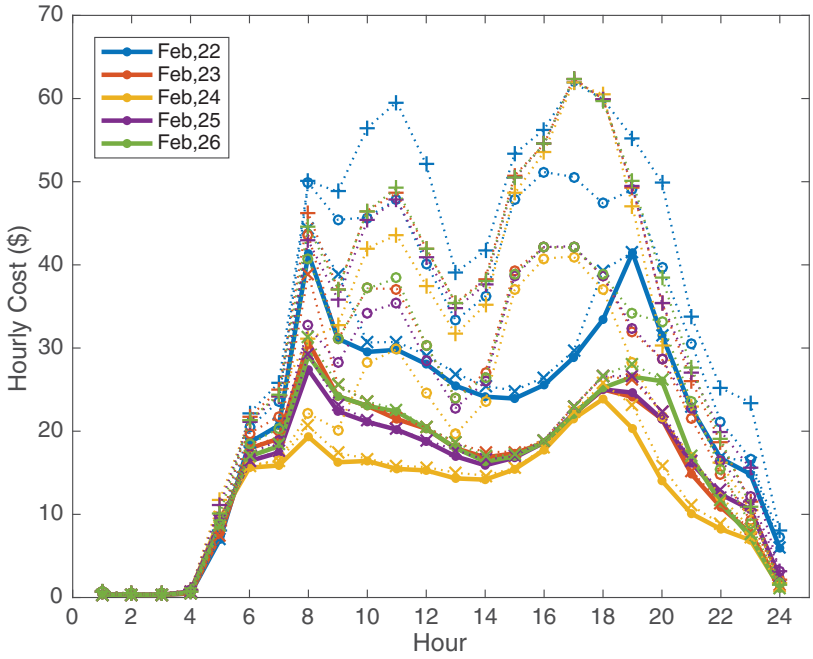


Figure 4.17: Comparison of the overall social costs of the 6 target buildings in the test winter week with different indoor temperature settings.

4.2 Case study of social energy implementation in a campus grid with dynamic building energy consumption model

In the above section Sec.4.1, we present an example of the technical virtual system for the DU's campus grid to demonstrate the concept of social energy. The technical virtual system receives and processes real world measurement data, evaluates target buildings' energy consumption, predicts energy suppliers' response, quantifies occupants' comfort level, optimizes between energy costs and occupants' efficiency costs, and eventually provides management advices for reducing the overall system costs. In order to better understand and observe the working mechanism of the technical virtual system, we simplified the buildings' energy consumption models, which assumes that consumption is determined solely by the current state and operation of the system. However in real situation, a building's energy consumption is not only related to ongoing "event" (ongoing operations, current environment parameters, current occupants' behaviors, *etc.*), but also affected by its previous status (consumption profiles from the passed period, previous environment parameters, operation strategy of facilities at previous time points, *etc.*).

Therefore, in this section, we intend to update the consumption model to a dynamic model and integrate it into the same socio-technical system developed above. Due to the limited access to the real system, we didn't add extra system parameters to describe the system's status, but sticking to the parameters and variables used in Sec.4.1. Let $\mathbf{x}_t$ denote the state of the entire system at time k, and it is composed of energy consumptions of the 6 target buildings $\mathbf{e}_t = [e_{1,t}\ e_{2,t} \cdots e_{6,t}]^\intercal$, outdoor temperature T_t^{out}, and hour of the day t, which is expressed in the following:

$$\mathbf{x}_t = [\mathbf{e}_t^\intercal,\ T_t^{out},\ t]^\intercal \tag{4.2.1}$$

And it follows the following state function

$$\mathbf{x}_{t+1} = \begin{bmatrix} \mathbf{e}_{t+1} \\ T_{t+1}^{out} \\ t+1 \end{bmatrix} = F(\mathbf{x}_t, \mathbf{u}_t) = \begin{bmatrix} H(\mathbf{e}_t, \mathbf{u}_t, T_{t+1}^{out}, t) \\ G(T_t^{out}) \\ t+1 \end{bmatrix}, \tag{4.2.2}$$

in which $\mathbf{e}_{t+1} = H_1(\mathbf{e}_t, \mathbf{u}_t, T_{t+1}^{out}, t)$ is the dynamic consumption model, and it is different from the consumption model defined in Equation 4.1.3. And $T_{t+1}^{out} = G(T_t^{out})$ is the outdoor temperature forecast model. Both of the models will be obtained by training time series nonlinear autoregressive with external input (NARX) Neural Networks.

Recall the cost function defined in Equations 4.1.12 - 4.1.14 in Section 4.1, in which the control variable $\mathbf{u_t}$ only affects the hourly cost $\psi_t = \Psi(\mathbf{u}_t)$, thus it is reasonable to set it as the objective function to be minimized. While in the case discussed in this section, where control decisions $\mathbf{u_t}$ influence all system states beyond the time point t, the goal of system management turns out to be minimizing accumulated cost beyond current time point t, no matter what management strategies have been applied before. To implement this idea, We first define the cost-to-go function $J(\cdot, \cdot)$ of the system as

$$J(\mathbf{x}_t, \underline{\mathbf{u}_t}) = \sum_{i=t}^{\infty} \gamma^{i-t} \cdot \Psi(\mathbf{x}_i, \mathbf{u}_i), \tag{4.2.3}$$

where $\Psi(\mathbf{x}_t, \mathbf{u}_t)$ is the utility function previously defined in Equations 4.1.12 - 4.1.14 which is recalled here as

$$\begin{aligned}\Psi(\mathbf{x}_t, \mathbf{u}_t) &= \mathbf{p}_t \cdot \mathbf{e}_t + \alpha(\mathbf{1} - \boldsymbol{\xi}_t) \cdot \mathbf{o}_t \\ &= \underbrace{\Gamma(\mathbf{u}_t, \mathbf{x}_t) \cdot H(\mathbf{u}_t, \mathbf{x}_t)}_{\text{energy cost}} + \underbrace{\alpha[1 - g(\mathbf{u}_t)] \cdot \mathbf{o}_t}_{\text{human cost}}. \end{aligned} \tag{4.2.4}$$

In the cost-to-go function 4.2.3, $\underline{\mathbf{u}_t} = (\mathbf{u}_t, \mathbf{u}_{t+1}, \mathbf{u}_{t+2}, \cdots)$ is the control sequence beyond time t, and $0 < \gamma < 1$ is the discount factor. And our goal is to find the optimal control sequence $\underline{\mathbf{u}_t^*} = (\mathbf{u}_t^*, \mathbf{u}_{t+1}^*, \mathbf{u}_{t+2}^*, \cdots)$, which results in the optimal (specifically minimum in this case study) cost-to-go $J^*(\mathbf{x}_t)$, which is

$$J^*(\mathbf{x}_t) = \inf_{\underline{\mathbf{u}_t}} J(\mathbf{x}_t, \underline{\mathbf{u}_t}) = J(\mathbf{x}_t, \underline{\mathbf{u}_t^*}). \tag{4.2.5}$$

It is obvious that Equation 4.2.5 can hardly be solved directly for highly nonlinear systems, especially in our case, the calculation of real-time electricity prices is achieved by solving another optimization problem, which means we don't have a deterministic function to describe the relationship between electricity prices and loads. And it is inferred that we are doing optimization problem with "internal black boxes". Therefore we need a more robust method and we propose to apply Neural Network based iterative adaptive dynamic programming algorithm [88, 91–93] to solve the minimum cost problem.

4.2.1 Neural Network based iterative adaptive dynamic programming for minimizing aggregated overall cost of the social energy system

According to Equations 4.2.3 and 4.2.5, we can rewrite the optimal cost-to-go function as

$$\begin{aligned} J^*(\mathbf{x}_t) &= \Psi(\mathbf{x}_t, \mathbf{u}_t^*) + \sum_{i=t+1}^{\infty} \gamma^{i-t} \Psi(\mathbf{x}_i, \mathbf{u}_i^*) \\ &= \Psi(\mathbf{x}_t, \mathbf{u}_t^*) + \gamma J^*(\mathbf{x}_{t+1}) \\ &= \min_{\mathbf{u}_t} \{\Psi(\mathbf{x}_t, \mathbf{u}_t) + \gamma J^*(\mathbf{x}_{t+1})\}. \end{aligned} \tag{4.2.6}$$

This is the Bellman optimality equation for dynamic programming. it indicates that, assuming the optimal control sequence from $t+1$ has been achieved and applied to the system, the optimal control vector for time t can be obtained by

$$\begin{aligned} \mathbf{u}_t^* &= \arg\min_{\mathbf{u}_t} \{\Psi(\mathbf{x}_t, \mathbf{u}_t) + \gamma J^*(\mathbf{x}_{t+1})\} \\ &= \arg\min_{\mathbf{u}_t} \{\Psi(\mathbf{x}_t, \mathbf{u}_t) + \gamma J^*(F(\mathbf{x}_t, \mathbf{u}_t)\}, \end{aligned} \tag{4.2.7}$$

where $\mathbf{x}_{t+1} = F(\mathbf{x}_t, \mathbf{u}_t)$ is the state function defined in 4.2.2. It allows us to optimize system performance over only one time step, which largely reduces the computational complexity and makes the solution to this problem feasible. Therefore, the major task here becomes finding the solution of the state dependent optimal cost-to-go function $J^*(\mathbf{x}_t)$.

First proposed by Paul J. Werbos in [91], and widely used in control theories later, the core of adaptive dynamic programming [94], also called approximate dynamic programming in other literatures [91, 93], is to design a critic network for approximate the cost-to-go function J through training. And iterative adaptive dynamic programming combined and integrated value iteration into the training

process of the critic network. The critic network updating mechanism is described in the diagram shown in Fig. 4.18.

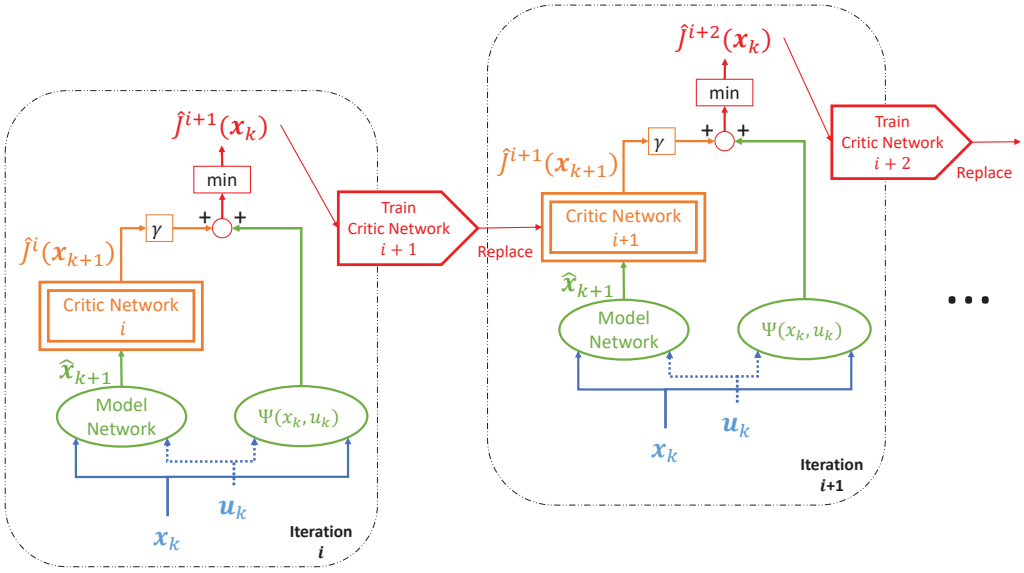


Figure 4.18: Update mechanism of the critic network in iterative adaptive dynamic programming.

As most of the traditional value iteration algorithm, we choose the quadratic form of the system states as the initial value function $J^0(\mathbf{x}_t) = \mathbf{x}_t^\intercal \mathbf{P}_0 \mathbf{x}_t$, where $\mathbf{P}_0$ is a positive-definite matrix, and we choose $\mathbf{P}_0 = \lambda \mathbf{I}_n$, λ is an arbitrary positive scalar, and $\mathbf{I}_n$ is an $n \times n$ unity matrix, with $n = 8$ is the length of the state vector in our case. We define $J^i(\mathbf{x}_t)$ as the value function in the ith iteration, the iteration will start as

$$\begin{aligned} J^1(\mathbf{x}_t) &= \min_{\mathbf{u}_t}\{\Psi(\mathbf{x}_t, \mathbf{u}_t) + \gamma J^0(\mathbf{x}_{t+1})\} \\ &= \min_{\mathbf{u}_t}\{\Psi(\mathbf{x}_t, \mathbf{u}_t) + \gamma J^0(F(\mathbf{x}_t, \mathbf{u}_t))\}. \end{aligned}$$

Then construct and initialize a neural network of a certain structure as the critic network, and we denote this critic network as $NN : \{\mathcal{W}_0, \mathcal{B}_0\}$, where $\mathcal{W}_0$ is the set of

its weights and $\mathcal{B}_0$ is the set of its biases. Using the values $J^1(\mathbf{x}_t)$ as training targets, we can update the network parameters as $\{\mathcal{W}_0, \mathcal{B}_0\} \rightarrow \{\mathcal{W}_1, \mathcal{B}_1\}$. Substituting the initial value function with the newly obtained $J^1(\cdot)$, and the algorithm will iterate as

$$\begin{aligned} J^i(\mathbf{x}_t) &= \min_{\mathbf{u}_t}\{\Psi(\mathbf{x}_t, \mathbf{u}_t) + \gamma J^{i-1}(\mathbf{x}_{t+1})\} \\ &= \min_{\mathbf{u}_t}\{\Psi(\mathbf{x}_t, \mathbf{u}_t) + \gamma J^{i-1}(F(\mathbf{x}_t, \mathbf{u}_t))\} \end{aligned} \tag{4.2.8}$$

for iteration index $i = 2, 3, \cdots$. The parameters of the critic network $\{\mathcal{W}_{i-1}, \mathcal{B}_{i-1}\} \rightarrow \{\mathcal{W}_i, \mathcal{B}_i\}$ will be trained and updated using $J^i(\mathbf{x}_t)$ as targets. It can be proved [95] that when i approaches infinity $i \rightarrow \infty$, value function approaches the optimal cost-to-go function $J^i(\mathbf{x}_t) \rightarrow J^*(\mathbf{x}_t)$.

4.2.2 Proof of convergence of the iterative value function to the optimal cost-to-go function

Convergence of the iterative value function [88, 95] is proved in the following.

Theorem 4.2.1. *Suppose that there exist constants* $0 < \rho < \infty$, $0 \leqslant \underline{\varepsilon} \leqslant 1$, *and* $1 \leqslant \overline{\varepsilon} \leqslant \infty$, *which satisfy* $0 \leqslant J^*(\mathbf{x}_{t+1}) \leqslant \rho\Psi(\mathbf{x}_t, \mathbf{u}_t)$, *and* $0 \leqslant \underline{\varepsilon} J^*(\mathbf{x}_t) \leqslant J^0(\mathbf{x}_t) \leqslant \overline{\varepsilon} J^*(\mathbf{x}_t)$, $\forall \mathbf{x}_t$, *and* $\forall \mathbf{u}_t$, *then the following inequalities hold uniformly:*

$$[1 + \frac{\underline{\varepsilon} - 1}{(1 + \rho^{-1})^i}]J^*(\mathbf{x}_t) \leqslant J^i(\mathbf{x}_t) \leqslant [1 + \frac{\overline{\varepsilon} - 1}{(1 + \rho^{-1})^i}]J^*(\mathbf{x}_t) \tag{4.2.9}$$

Therefore, the value function $J^i(\mathbf{x}_t)$ *converges to* $J^*(\mathbf{x}_t)$ *when* i *approaches infinity, which is* $\lim_{i \rightarrow \infty} J^i(\mathbf{x}_t) = J^*(\mathbf{x}_t)$.

Proof. We first start from the lower bound of the inequality in Equation 4.2.9, when $i = 1$,

$$
\begin{aligned}
J^1(\mathbf{x}_t) &= \min_{\mathbf{u}_t}\{\Psi(\mathbf{x}_t,\mathbf{u}_t)+\gamma\cdot J^0(\mathbf{x}_{t+1})\} \\
&\geqslant \min_{\mathbf{u}_t}\{\Psi(\mathbf{x}_t,\mathbf{u}_t)+\gamma\cdot \underline{\varepsilon} J^*(\mathbf{x}_{t+1})\}. \qquad (4.2.10)
\end{aligned}
$$

If there exists a constant K_1, which satisfies $K_1 \leqslant 0$ and $1+\rho K_1 = \underline{\varepsilon} - K_1$, with $\rho\Psi(\mathbf{x}_t,\mathbf{u}_t) - \gamma\cdot J^*(\mathbf{x}_{t+1}) \geqslant 0$, Equation 4.2.10 can be recast as

$$
\begin{aligned}
J^1(\mathbf{x}_t) &\geqslant \min_{\mathbf{u}_t}\{\Psi(\mathbf{x}_t,\mathbf{u}_t)+\gamma\cdot\underline{\varepsilon}J^*(\mathbf{x}_{t+1})+K_1[\rho\Psi(\mathbf{x}_t,\mathbf{u}_t)-\gamma\cdot J^*(\mathbf{x}_{t+1})]\} \\
&= (1+\rho K_1)\cdot\min_{\mathbf{u}_t}\{\Psi(\mathbf{x}_t,\mathbf{u}_t)+\gamma\cdot J^*(\mathbf{x}_{t+1})\} \\
&= (1+\rho K_1)\cdot J^*(\mathbf{x}_t) \qquad (4.2.11)
\end{aligned}
$$

where $K_1 = \frac{\underline{\varepsilon}-1}{1+\rho} \leqslant 0$. Further when $i=2$, it can be deduced that

$$
\begin{aligned}
J^2(\mathbf{x}_t) &= \min_{\mathbf{u}_t}\{\Psi(\mathbf{x}_t,\mathbf{u}_t)+\gamma\cdot J^1(\mathbf{x}_{t+1})\} \\
&\geqslant \min_{\mathbf{u}_t}\{\Psi(\mathbf{x}_t,\mathbf{u}_t)+\gamma\cdot(1+\rho K_1)J^*(\mathbf{x}_{t+1})\} \\
&\geqslant \min_{\mathbf{u}_t}\{\Psi(\mathbf{x}_t,\mathbf{u}_t)+\gamma\cdot(1+\rho K_1)J^*(\mathbf{x}_{t+1})+K_2[\rho\Psi(\mathbf{x}_t,\mathbf{u}_t)-\gamma\cdot J^*(\mathbf{x}_{t+1})]\} \\
&= (1+\rho K_2)\cdot\min_{\mathbf{u}_t}\{\Psi(\mathbf{x}_t,\mathbf{u}_t)+\gamma\cdot J^*(\mathbf{x}_{t+1})\} \\
&= (1+\rho K_2)\cdot J^*(\mathbf{x}_t) \qquad (4.2.12)
\end{aligned}
$$

with $K_2 = \frac{1}{1+\rho^{-1}}\cdot K_1 \leqslant 0$ to satisfy $K_2 \leqslant 0$ and $1+\rho K_2 = (1+\rho K_1) - K_2$. And further when $i=i$, the following can be easily obtained,

$$
\begin{aligned}
J^i(\mathbf{x}_t) &= \min_{\mathbf{u}_t}\{\Psi(\mathbf{x}_t,\mathbf{u}_t)+\gamma\cdot J^{i-1}(\mathbf{x}_{t+1})\} \\
&\geqslant \min_{\mathbf{u}_t}\{\Psi(\mathbf{x}_t,\mathbf{u}_t)+\gamma\cdot(1+\rho K_{i-1})J^*(\mathbf{x}_{t+1})\} \\
&\geqslant \min_{\mathbf{u}_t}\{\Psi(\mathbf{x}_t,\mathbf{u}_t)+\gamma\cdot(1+\rho K_{i-1})J^*(\mathbf{x}_{t+1})+K_i[\rho\Psi(\mathbf{x}_t,\mathbf{u}_t)-\gamma\cdot J^*(\mathbf{x}_{t+1})]\} \\
&= (1+\rho K_i)\cdot\min_{\mathbf{u}_t}\{\Psi(\mathbf{x}_t,\mathbf{u}_t)+\gamma\cdot J^*(\mathbf{x}_{t+1})\} \\
&= (1+\rho K_i)\cdot J^*(\mathbf{x}_t), \qquad (4.2.13)
\end{aligned}
$$

where $K_i=\frac{1}{1+\rho^{-1}}\cdot K_{i-1}\leqslant 0$. And by mathematical induction we will have

$$
K_i=\frac{1}{(1+\rho^{-1})^{i-1}}\cdot K_1=\frac{\underline{\varepsilon}-1}{(1+\rho^{-1})^i}\rho^{-1} \qquad (4.2.14)
$$

$$
J^i(\mathbf{x}_t)\geqslant[1+\frac{\underline{\varepsilon}-1}{(1+\rho^{-1})^i}]\cdot J^*(\mathbf{x}_t) \qquad (4.2.15)
$$

Therefore, the lower bound of the inequality 4.2.9 is proved. Following the same procedures, the upper bound of the inequality can be proved. □

4.2.3 Procedures to train the critic network for approximating the optimal cost-to-to function

Since the rationale and feasibility of finding the optimal cost-to-go function for solving the minimum system cost problem have been explained and proved, here we summarize the detailed step-by-step implementation of the iterative adaptive dynamic programming algorithm.

(i) Train NARX consumption model $\mathbf{e}_{t+1}=H(\mathbf{e}_t,\mathbf{u}_t,T^{out}_{t+1},t)$ and outdoor temperature forecast model $T^{out}_{t+1}=G(T^{out}_t)$. And form the system state model $\mathbf{x}_{t+1}=F(\mathbf{x}_t,\mathbf{u}_t)$.

(ii) Set the value of max iteration i_{max}, and desired accuracy ϑ for the critic network training.

(iii) Randomly choose ι state vectors $\{\mathbf{x}_t^1, \mathbf{x}_t^2, \cdots, \mathbf{x}_t^\iota\}$. Set the discount factor $\gamma = 0.6$ in the case study. And construct and initialize the critic network's parameters $\{\mathcal{W}_0, \mathcal{B}_0\}$.

(iv) Set the iteration index $i = 0$, and set $\mathbf{P}_0 = \lambda \cdot \mathbf{I}_8$.

(v) Set the iteration index $i = i+1$, and solve Equation 4.2.8 to obtain the updated estimation of the critic network $\{J^i(\mathbf{x}_t^1), J^i(\mathbf{x}_t^2), \cdots, J^i(\mathbf{x}_t^\iota)\}$.

(vi) Train the critic network and update its parameters $\{\mathcal{W}_i, \mathcal{B}_i\}$.

(vii) If $i > 1$, go to (viii); Else if $J^0(\mathbf{x}_t^\iota) > J^1(\mathbf{x}_t^\iota), \forall \mathbf{x}_t^\iota$, go to (viii); Otherwise, increase λ and go back to (iv).

(viii) If $i > i_{max}$, or $|J^i(\mathbf{x}_t^j) - J^j(\mathbf{x}_t^\iota)| \leqslant \vartheta$, for $j = 1, 2, \cdots, \iota$, go to (ix); Otherwise, go back to (v).

(ix) The optimal cost-to-go function is achieved by $J^*(\mathbf{x}_t) = J^i(\mathbf{x}_t)$, and the neural network $NN : \{\mathcal{W}_i, \mathcal{B}_i\}$ can output for further computation.

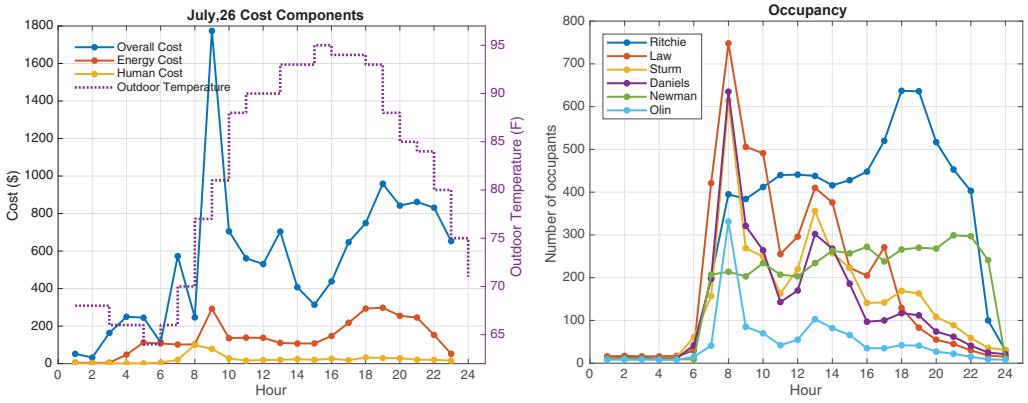
(x) Terminate the iterative algorithm.

In the numerical experiment conducted in this case study, we set and $\iota = 240$ and $\gamma = 0.6$. Constant $\lambda = 1$ was initially set, but updated by the algorithm to $\lambda = 5$. And the algorithm eventually stopped at $i = 103$.

4.2.4 Result and discussion

Same as in Section 4.1, we chose the same days (5 successive days in summer, and 5 successive days in winter), to demonstrate and test the concept and algorithm in

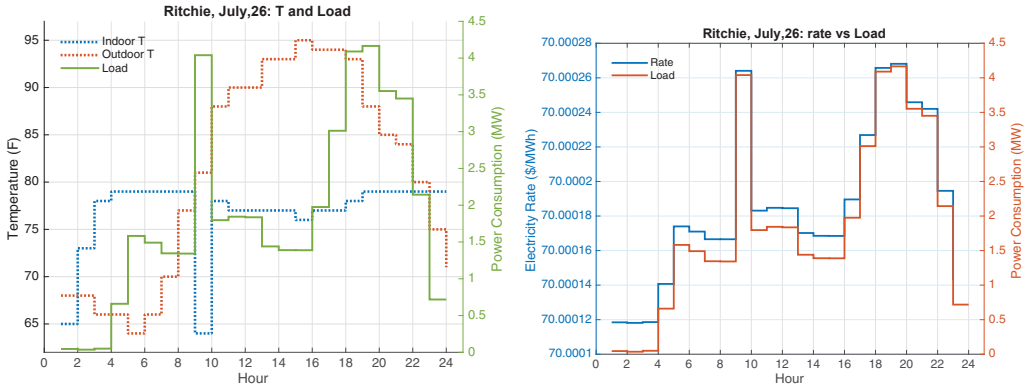
this section. However, since the amount of figures generated is huge, we only selected test results from one of these days, July 26th, to present here.



(a) (b)

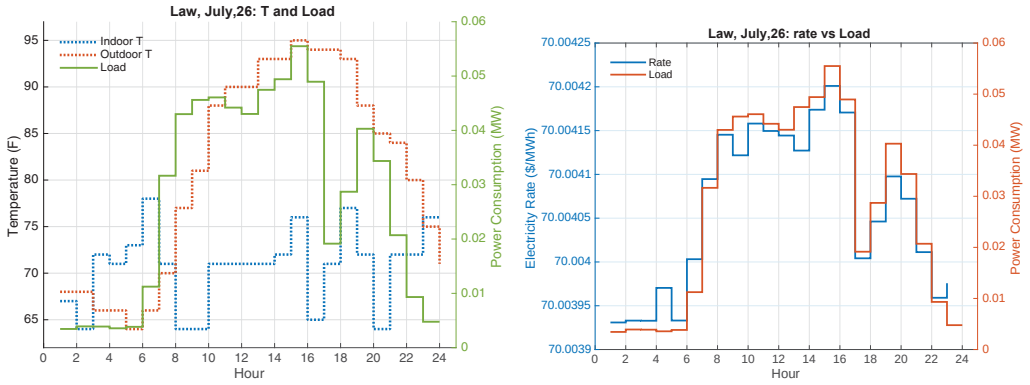
Figure 4.19: (a) Outdoor temperatures and the cost components for the entire system, (b) Number of Occupants in the 6 target buildings.

Figure 4.19(a) plots out the optimized aggregated overall cost for each hour of the testing day, and the corresponding energy cost and human working efficiency cost are also provided. As part of the system's state, as well as the driven variables of the campus energy system, outdoor temperature changes and number of occupants in the 6 target buildings are shown in Figure 4.19. And Figures 4.20 to 4.25 look at the 6 target buildings separately by presenting the individual HVAC management strategies (indoor temperature control), energy consumption profiles and hourly real-time electricity prices information for each of them.



(a) (b)

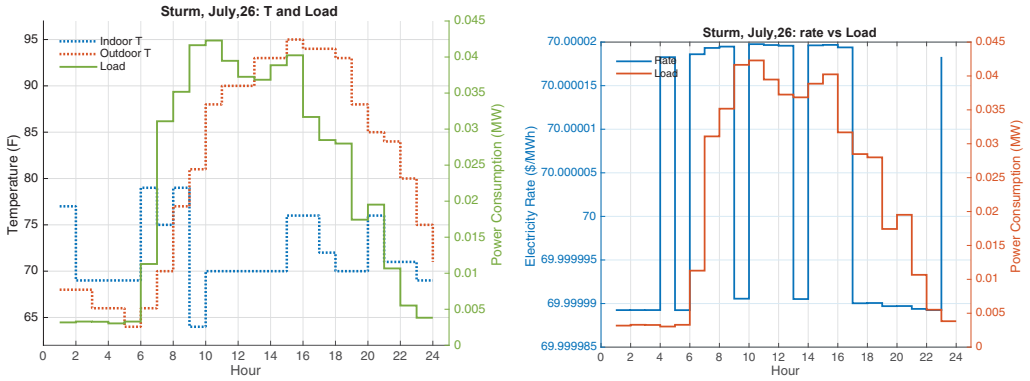
Figure 4.20: (a) Suggested indoor temperature settings vs. outdoor temperatures, and the power consumption for Ritchie Center, (b) Real-time hourly electricity rate vs. power consumption for Ritchie Center.



(a) (b)

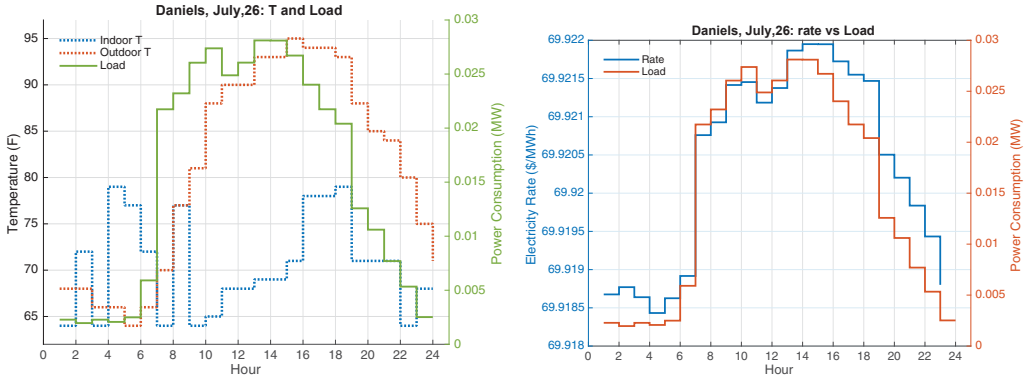
Figure 4.21: (a) Suggested indoor temperature settings vs. outdoor temperatures, and the power consumption for the Law building, (b) Real-time hourly electricity rate vs. power consumption for the Law building.

We can see from the overall costs plotted in Figure 4.19(a), that it presents not only current system cost, but also an estimation of future aggregated costs applying the current management strategy. By comparing the consumption profiles in Figures 4.20 to 4.25, it is obvious that Ritchie Center's energy consumption is much higher than others', making it the major contributor to the energy cost in the system.



(a) (b)

Figure 4.22: (a) Suggested indoor temperature settings vs. outdoor temperatures, and the power consumption for Sturm Hall, (b) Real-time hourly electricity rate vs. power consumption for Sturm Hall.

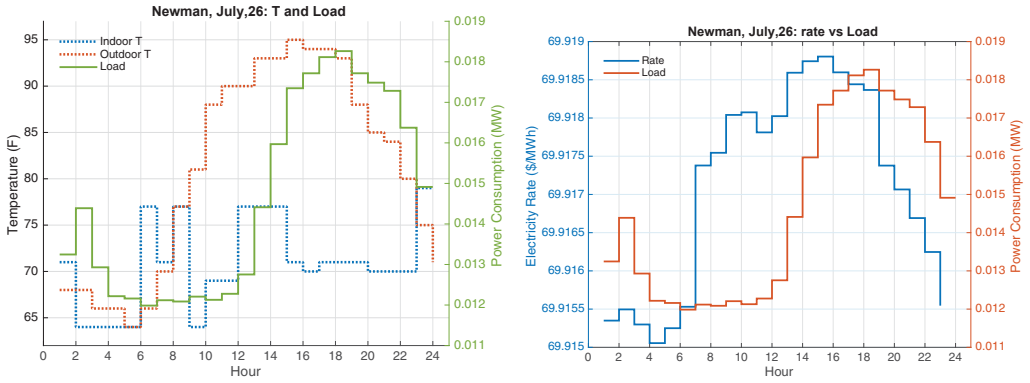


(a) (b)

Figure 4.23: (a) Suggested indoor temperature settings vs. outdoor temperatures, and the power consumption for the Daniels buildings, (b) Real-time hourly electricity rate vs. power consumption for the Daniels buildings.

And let's look at two featured point during the test period. The first one is $9:00$ o'clock, when the working hour begins and a sharp increase in amount of occupants appears. Due to this change, a peak of all types of costs occurs. And at this time the total amount of occupants in the system reflects to be the largest, which leads to a trend for most of the academic buildings to set their indoor temperatures to $71-72°F$, which induces the highest working efficiency and in turn guaranties

the lowest human costs. And at this time, Ritchie Center's temperature setting drops to a low value. Although this setting seems to be unwise, for it leads to high consumption for that hour, but it is the one suggested by the evaluation algorithm as it may lay good start for the future times.



(a) (b)

Figure 4.24: (a) Suggested indoor temperature settings vs. outdoor temperatures, and the power consumption for Newman Center, (b) Real-time hourly electricity rate vs. power consumption for Newman Center.

The second featured point is 18 : 00 o'clock. At this time point, the outdoor environment still remains hot, occupants of the academic buildings start to leave, while visitor amount of Ritchie Center and Newman Center maintains large. Here, we can observe a very random selection of indoor temperatures in those academic buildings, since human cost is no longer the leading factor at this time.

4.3 Chapter Conclusion

This chapter presents an initial investigation into the intension and extension of the social energy concept. A case study of consumer level social energy is provided in the context of DU campus power grid, involving power system simulation, numerical experiments for smart building modeling, distribution level real-time pricing and social response to the pricing signals. Although the energy system studied in this

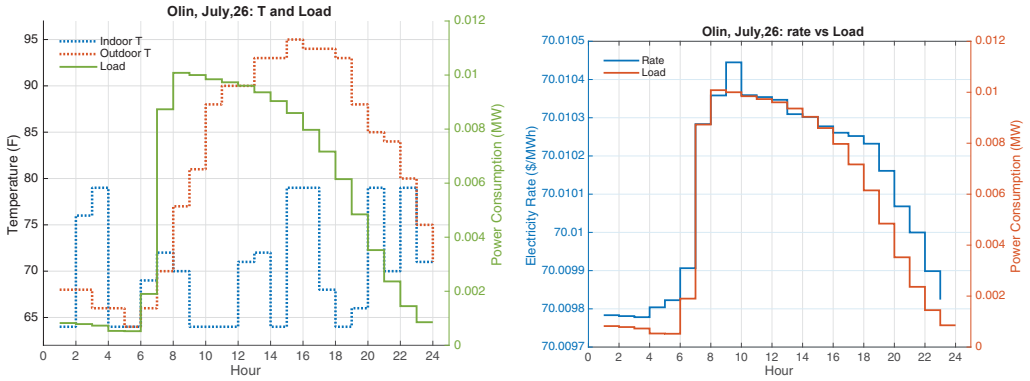


Figure 4.25: (a) Suggested indoor temperature settings vs. outdoor temperatures, and the power consumption for Olin Hall, (b) Real-time hourly electricity rate vs. power consumption for Olin Hall.

chapter is largely simplified, it still gives us a vision into the interactions among the users, energy systems and the electricity suppliers.

What's more, the Neural Network based iterative adaptive dynamic programming algorithm was introduced for solving highly nonlinear cost minimization problem with "black-box" models included. Convergence of the solution and detailed neural network implementation procedures are provided. This robust algorithm can reduce the complexity of such computation consuming problems, and reveals high feasibility and ability for a solution even when the case approaches real world systems.

The case study only provides an application scenario of a relatively small smart community, but our vision for social energy is that, it can be applicable in different social and technical scales and provide a multi-layered solution to benefit the socio-technical system.

Chapter 5

Conclusion

The request for big data based situation awareness is urgent today. In this dissertation, two data-centric approaches for power system situation awareness and management are proposed to address the security problems in the transmission and/or distribution grid and social benefits augmentation problem at the distribution - customer lever, respectively.

The first approach provides a fault analysis solution based on characterization and analytics of the synchrophasor measurements. Specifically, the OSMDSA and MPD based spatial-temporal synchrophasor data characterization method was developed to reduce data volume while preserving comprehensive information for the big data analyses. And the WGC method was investigated to conduct fault impact causal analysis during system disturbance for fault localization. Numerical results and comparison with other methods demonstrate the effectiveness and robustness of this analytics approach.

The second approach investigates the concept and application of social energy upon the University of Denver campus grid to provide management improvement solutions for optimizing social cost. Social element – human working productivity cost, and economic element – electricity consumption cost, are included in the eval-

uation of social cost. Moreover, power system simulation, numerical experiments for smart building modeling, distribution level real-time pricing and social response to the pricing signals are studied for implementing the interactive artificial-physical management scheme. What's more, the robust iterative adaptive dynamic programming algorithm was introduced, which is capable of solving complex system control problems.

The following directions are proposed as possible future works for the research work done in this dissertation.

- Current research work only includes 6 buildings in a small scale power grid as investigation object, and the technical system structure is simplified, which considers the HVAC system with indoor temperature control as the control means. To expand this work for generating more realistic benefits to society, updating current system to approach real world system is a promising direction. As more elements and factors, which affect the system's performance in various ways, robustness of the proposed algorithm can also be examined.

- As the interface to integrate social elements into the socio-technical system, development of more advanced human behavior and feedback modeling could be conducted.

- The demand-side management scheme investigated in this dissertation is a centralized scheme, which is usually time consuming. With the large scale applications of decentralized algorithms, comparison research using decentralized management algorithms, such as game theory, or development of centralized-distributed combined schemes could be conducted.

Bibliography

[1] Leonard L. Grigsby, Ed., *The Electric Power Engineering Handbook: Power System Stability and Control, Third Edition*, CRC Press, 2012.

[2] "U.S. - Canada power system outage task force: Final report on the implementation of task force recommendations," Tech. Rep., Office of Electricity Delivery & Energy Reliability, 2004.

[3] Mladen Kezunovic, Le Xie, and Santiago Grijalva, "The role of big data in improving power system operation and protection," *2013 IREP Symposium-Bulk Power System Dynamics and Control*, 2013.

[4] Huaiguang Jiang, Yingchen Zhang, Jun Jason Zhang, and Muljadi Eduard, "PMU-aided voltage security assessment for a wind power plant," *2015 IEEE Power & Energy Society General Meeting*, pp. 1 – 5, 2015.

[5] Yunfeng Cheng, Chao Lu, Kun Men, and Liang Tu, "Research on perception of power system state based on wams," in *2015 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, 2015, pp. 1–5.

[6] Anjan Bose, "Smart transmission grid applications and their supporting infrastructure," *IEEE Transcations on Smart Grid*, vol. 1, pp. 11–19, 2010.

[7] Jun Xiao, Xiao-dan Guo, and Lin-quan Bai, "An improved model of total supply capability for distribution systems," in *Power and Energy Engineering Conference (APPEEC), 2012 Asia-Pacific*. IEEE, 2012, pp. 1–4.

[8] M. K. Banavar, J. Zhang, B. Chakraborty, H. Kwon, Y. Li, H. Jiang, A. Spanias, C. Tepedelenlioglu, C. Chakrabarti, and A. Papandreou-Suppappola, "An overview of recent advances on distributed and agile sensing algorithms and implementation," *Digital Signal Processing*, vol. 39, pp. 1–14, April 2015.

[9] Min Chen, Shiwen Mao, and Yunhao Liu, "Big data: A survey," *Mobile Networks and Applications*, vol. 19, no. 2, pp. 171–209, 2014.

[10] Yi Gu, Huaiguang Jiang, Yingchen Zhang, and David Wenzhong Gao, "Statistical scheduling of economic dispatch and energy reserves of hybrid power systems with high renewable energy penetration," in *2014 48th Asilomar Conference on Signals, Systems and Computers*, 2014, pp. 530–534.

[11] Yogesh Simmhan, Saima Aman, Alok Kumbhare, Rongyang Liu, Sam Stevens, Qunzhi Zhou, and Viktor Prasanna, "Cloud-based software platform for big data analytics in smart grids," *Computing in Science & Engineering*, vol. 15, no. 4, pp. 38–47, 2013.

[12] Sebnem Rusitschka, Kolja Eger, and Christoph Gerdes, "Smart grid data cloud: A model for utilizing cloud computing in the smart grid domain," in *2010 First IEEE International Conference on Smart Grid Communications (SmartGridComm)*. IEEE, 2010, pp. 483–488.

[13] Farrokh Aminifar, Amin Khodaei, Mahmud Fotuhi-Firuzabad, and Mohammad Shahidehpour, "Contingency-constrained pmu placement in power networks," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 516–523, 2010.

[14] Farrokh Aminifar, Caro Lucas, Amin Khodaei, and Mahmud Fotuhi-Firuzabad, "Optimal placement of phasor measurement units using immunity genetic algorithm," *IEEE Transactions on Power Delivery*, vol. 24, no. 3, pp. 1014–1020, 2009.

[15] S.G Mallat and Z Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, pp. 3397–3415, 1993.

[16] Scott Shaobing Chen, David L Donoho, and Michael A Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.

[17] Huaiguang Jiang, Jun Jason Zhang, Adam Hebb, and Mohammad H Mahoor, "Time-frequency analysis of brain electrical signals for behvior recognition in patients with parkinson's disease," in *2013 Asilomar Conference on Signals, Systems and Computers*. IEEE, 2013, pp. 1843–1847.

[18] JP Paul, JY Leost, and JM Tesseron, "Survey of the secondary voltage control in france: present realization and investigations," *IEEE Transactions on Power Systems*, vol. 2, no. 2, pp. 505–511, 1987.

[19] P Lagonotte, JC Sabonnadiere, JY Leost, and JP Paul, "Structural analysis of the electrical system: application to secondary voltage control in france," *IEEE Transactions on Power Systems*, vol. 4, no. 2, pp. 479–486, 1989.

[20] Bei Gou, "Generalized integer linear programming formulation for optimal pmu placement," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1099–1104, 2008.

[21] Huaiguang Jiang, Jun J Zhang, Wenzhong Gao, and Ziping Wu, "Fault detection, identification, and location in smart grid based on data-driven computa-

tional methods," *IEEE Transactions on Smart Grid*, vol. 5, pp. 2947 – 2956, 2014.

[22] Raj Aggarwal and Yonghua Song, "Artificial neural networks in power systems. III. examples of applications in power systems," *Power Engineering*, vol. 12, pp. 279–287, 1998.

[23] Peter G. V. Axelberg, Irene Yu-Hua Gu, and Math H. J. Bollen, "Support vector machine for classification of voltage disturbances," *IEEE Transactions on Power Delivery*, vol. 22, pp. 1297–1303, 2007.

[24] L. Barnett and A. K. Seth, "The MVGC multivariate granger causality toolbox: A new approach to granger-causal inference," *Journal of neuroscience methods*, vol. 223, pp. 50–68, 2014.

[25] Jacopo Torriti, *Peak energy demand and demand side response*, Routledge, 2015.

[26] VSK Murthy Balijepalli, Vedanta Pradhan, SA Khaparde, and RM Shereef, "Review of demand response under smart grid paradigm," in *2011 IEEE PES on Innovative Smart Grid Technologies-India (ISGT India)*. IEEE, 2011, pp. 236–243.

[27] J. S. Vardakas, N. Zorba, and C. V. Verikoukis, "A survey on demand response programs in smart grids: Pricing methods and optimization algorithms," *IEEE Communications Surveys Tutorials*, vol. 17, no. 1, pp. 152–178, March 2015.

[28] Michael Angelo A Pedrasa, Ted D Spooner, and Iain F MacGill, "Coordinated scheduling of residential distributed energy resources to optimize smart home energy services," *IEEE Transactions on Smart Grid*, vol. 1, no. 2, pp. 134–143, 2010.

[29] Eunji Lee and Hyokyung Bahn, "Electricity usage scheduling in smart building environments using smart devices," *The Scientific World Journal*, vol. 2013, 2013.

[30] A-H Mohsenian-Rad, Vincent WS Wong, Juri Jatskevich, and Robert Schober, "Optimal and autonomous incentive-based energy consumption scheduling algorithm for smart grid," in *IEEE Transactions on Innovative Smart Grid Technologies (ISGT)*, 2010, pp. 1–6.

[31] Nikolaos Gatsis and Georgios B Giannakis, "Cooperative multi-residence demand response scheduling," in *2011 45th Annual Conference on Information Sciences and Systems (CISS)*. IEEE, 2011, pp. 1–6.

[32] "Monthly Energy Review: July 2017 version," Tech. Rep., U.S. Energy Information Administration, 2017.

[33] Joseph S Sussman, *Perspectives on intelligent transportation systems (ITS)*, Springer Science & Business Media, 2008.

[34] Michael D Meyer, "Demand management as an element of transportation policy: using carrots and sticks to influence travel behavior," *Transportation Research Part A: Policy and Practice*, vol. 33, no. 7, pp. 575–599, 1999.

[35] Salman Mohagheghi and Neda Raji, "Intelligent demand response scheme for energy management of industrial systems," in *IEEE Industry Applications Society Annual Meeting (IAS),*. IEEE, 2012, pp. 1–9.

[36] Daniel Olsen, "Opportunities for energy efficiency and demand response in the california cement industry," 2012.

[37] Arne Grein and Martin Pehnt, "Load management for refrigeration systems: Potentials and barriers," *Energy Policy*, vol. 39, no. 9, pp. 5598–5608, 2011.

[38] Johanna L Mathieu, Phillip N Price, Sila Kiliccote, and Mary Ann Piette, "Quantifying changes in building electricity use, with application to demand response," *IEEE Transactions on Smart Grid*, vol. 2, no. 3, pp. 507–518, 2011.

[39] Naoya Motegi, Mary Ann Piette, David S Watson, Sila Kiliccote, and Peng Xu, "Introduction to commercial building control strategies and techniques for demand response," *Lawrence Berkeley National Laboratory LBNL-59975*, 2007.

[40] Sila Kiliccote, "Findings from seven years of field performance data for automated demand response in commercial buildings," *Lawrence Berkeley National Laboratory*, 2010.

[41] Mehdi Maasoumy, Catherine Rosenberg, Alberto Sangiovanni-Vincentelli, and Duncan S Callaway, "Model predictive control approach to online computation of demand-side flexibility of commercial buildings hvac systems for supply following," in *American Control Conference (ACC)*. IEEE, 2014, pp. 1082–1089.

[42] Tao Zhang, Peer-Olaf Siebers, and Uwe Aickelin, "Modelling electricity consumption in office buildings: An agent based approach," *Energy and Buildings*, vol. 43, no. 10, pp. 2882–2892, 2011.

[43] Mehreen S Gul and Sandhya Patidar, "Understanding the energy consumption and occupancy of a multi-purpose academic building," *Energy and Buildings*, vol. 87, pp. 155–165, 2015.

[44] T. Logenthiran, D. Srinivasan, and T. Z. Shun, "Demand side management in smart grid using heuristic optimization," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1244–1252, Sept 2012.

[45] Faisal A Mohamed and Heikki N Koivo, "System modelling and online optimal management of microgrid using mesh adaptive direct search," *International*

Journal of Electrical Power & Energy Systems, vol. 32, no. 5, pp. 398–407, 2010.

[46] Jamshid Aghaei and Mohammad-Iman Alizadeh, "Multi-objective self-scheduling of CHP (combined heat and power)-based microgrids considering demand response programs and ESSs (energy storage systems)," *Energy*, vol. 55, pp. 1044–1054, 2013.

[47] Di Zhang, Nilay Shah, and Lazaros G Papageorgiou, "Efficient energy consumption and operation management in a smart building with microgrid," *Energy Conversion and Management*, vol. 74, pp. 209–222, 2013.

[48] Jidong Wang, Zhiqing Sun, Yue Zhou, and Jiaqiang Dai, "Optimal dispatching model of smart home energy management system," in *2012 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia)*. IEEE, 2012, pp. 1–5.

[49] Marco Zugno, Juan Miguel Morales, Pierre Pinson, and Henrik Madsen, "A bilevel model for electricity retailers' participation in a demand response market environment," *Energy Economics*, vol. 36, pp. 182–197, 2013.

[50] Claes Sandels, Daniel Brodén, Joakim Widén, Lars Nordström, and Enar Andersson, "Modeling office building consumer load with a combined physical and behavioral approach: Simulation and validation," *Applied Energy*, vol. 162, pp. 472–485, 2016.

[51] JM Yusta, HM Khodr, and AJ Urdaneta, "Optimal pricing of default customers in electrical distribution systems: Effect behavior performance of demand response models," *Electric Power Systems Research*, vol. 77, no. 5, pp. 548–558, 2007.

[52] Albert Molderink, Vincent Bakker, Maurice GC Bosman, Johann L Hurink, and Gerard JM Smit, "Domestic energy management methodology for optimizing

efficiency in smart grids," in *2009 IEEE Bucharest on PowerTech,*, 2009, pp. 1–7.

[53] Liang Zhou, Joel JPC Rodrigues, and Luís M Oliveira, "Qoe-driven power scheduling in smart grid: architecture, strategy, and methodology," *IEEE Communications Magazine*, vol. 50, no. 5, pp. 136–141, 2012.

[54] Jin Xiao, Jian Li, Raouf Boutaba, and James Won-Ki Hong, "Comfort-aware home energy management under market-based demand-response," in *Proceedings of the 8th International Conference on Network and Service Management*. International Federation for Information Processing, 2012, pp. 10–18.

[55] Pengwei Du and Ning Lu, "Appliance commitment for household load scheduling," *IEEE transactions on Smart Grid*, vol. 2, no. 2, pp. 411–419, 2011.

[56] Olli Seppanen, William J Fisk, and Q. Lei, "Effect of temperature on task performance in office environment," *Lawrence Berkeley National Laboratory*, 2006.

[57] Olli Seppanen, William J Fisk, and David Faulkner, "Control of temperature for health and productivity in offices," *Lawrence Berkeley National Laboratory*, 2004.

[58] RS Ferreira, LA Barroso, and MM Carvalho, "Demand response models with correlated price data: A robust optimization approach," *Applied energy*, vol. 96, pp. 133–149, 2012.

[59] D. O'Neill, M. Levorato, A. Goldsmith, and U. Mitra, "Residential demand response using reinforcement learning," in *2010 First IEEE International Conference on Smart Grid Communications*, Oct 2010, pp. 409–414.

[60] Nikolaos Gatsis and Georgios B Giannakis, "Residential demand response with interruptible tasks: Duality and algorithms," in *2011 50th IEEE Conference*

on Decision and Control and European Control Conference (CDC-ECC). IEEE, 2011, pp. 1–6.

[61] Stephen Boyd and Lieven Vandenberghe, *Convex optimization*, Cambridge university press, 2004.

[62] Michael G Kallitsis, George Michailidis, and Michael Devetsikiotis, "Optimal power allocation under communication network externalities," *IEEE Transactions on Smart Grid*, vol. 3, no. 1, pp. 162–173, 2012.

[63] SO Orero and MR Irving, "A genetic algorithm for network partitioning in power system state estimation," in *UKACC International Conference on Control'96,*, 1996, vol. 1, pp. 162–165.

[64] M Ilic'-Spong, J Christensen, and KL Eichorn, "Secondary voltage control using pilot point information," *IEEE Transactions on Power Systems*, vol. 3, no. 2, pp. 660–668, 1988.

[65] J. L. Sancha, J. L. Fernandez, A. Cortes, and J. Abarca, "Secondary voltage control: analysis, solutions and simulation results for the spanish transmission system," *IEEE Transactions on Power Systems*, vol. 11, pp. 630 – 638, 1996.

[66] Prabha Kundur, *Power System Stability and Control*, McGraw-Hill Professional, 1994.

[67] G Yesuratnam and D Thukaram, "Congestion management in open access based on relative electrical distances using voltage stability criteria," *Electric power systems research*, vol. 77, no. 12, pp. 1608–1618, 2007.

[68] K Visakha, D Thukaram, and Lawrence Jenkins, "Transmission charges of power contracts based on relative electrical distances in open access," *Electric Power Systems Research*, vol. 70, no. 2, pp. 153–161, 2004.

[69] J.B. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, 1967.

[70] D.J.C. MacKay, *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press, 2003.

[71] Huaiguang Jiang, Lei Huang, Jun Jason Zhang, Yingchen Zhang, and David Wenzhong Gao, "Spatial-temporal characterization of synchrophasor measurement systems-a big data approach for smart grid system situational awareness," in *IEEE 48th Asilomar Conference on Signals, Systems and Computers*, 2014, pp. 750–754.

[72] Lei Huang, Yuanzhang Sun, Jian Xu, Wenzhong Gao, Jun Zhang, and Ziping Wu, "Optimal PMU placement considering controlled islanding of power system," *IEEE Transactions on Power Systems*, vol. 29, pp. 742–755, 2014.

[73] Chunchun Xu, Zhian Zhong, Virgilio Centeno, Richard Conners, and Yilu Liu, "Practical issues in frequency disturbance recorder design for wide-area monitoring," *Electrial Power Quality and Utilisation*, vol. 11, pp. 69–76, 2005.

[74] Zhian Zhong, Chunchun Xu, Bruce J. Billian, Li Zhang, Shu-Jen Steven Tsai, Virgilio A. Centeno Richard W. Conners, Arun G. Phadke, and Yilu Liu, "Power system frequency monitoring network (fnet) implementation," *IEEE Transactions on Power Systems*, vol. 20, pp. 1914–1921, 2005.

[75] *A tutorial on hidden Markov models and selected applications in speech recognition*, 1989.

[76] Anil K Seth, "A matlab toolbox for granger causal connectivity analysis," *Journal of neuroscience methods*, vol. 186, no. 2, pp. 262–273, 2010.

[77] Sébastien David, Miguel A Ferrer, Carlos M Travieso, Jesús B Alonso, and Dpto De Señales y Comunicaciones, "gpdshmm: A hidden markov model toolbox in the matlab environment," *Complex Systems Intelligence and Modern Technological Applications*, pp. 476–479, 2004.

[78] Bei Gou, "Generalized integer linear programming formulation for optimal PMU placement," *IEEE Transactions on Power Systems*, vol. 23, 2008.

[79] Mohammad Reza Aghamohammadi and Ali Shahmohammadi, "Intentional islanding using a new algorithm based on ant search mechanism," *International Journal of Electrical Power & Energy Systems*, vol. 35, 2012.

[80] Po-Chen Chen, Vuk Malbasa, Yongsheng Dong, and Mladen Kezunovic, "Sensitivity analysis of voltage sag based fault location with distributed generation," *IEEE Transactions on Smart Grid*, vol. 6, pp. 2098 – 2106, 2015.

[81] Po-Chen Chen, Vuk Malbasa, and Mladen Kezunovic, "Sensitivity analysis of voltage sag based fault location algorithm," in *IEEE 2014 Power Systems Computation Conference (PSCC)*, 2014, pp. 1–7.

[82] Mladen Kezunovic, "Smart fault location for smart grids," *IEEE Transactions on Smart Grid*, vol. 2, no. 1, pp. 11–22, 2011.

[83] Steve Tom, "Managing energy and comfort," *ASHRAE Journal*, vol. 50, no. 6, pp. 18–27, 2008.

[84] S.S. Lang, "Study links warm offices to fewer typing errors and higher productivity," Cornell Chronicle, Oct. 2004, http://www.news.cornell.edu/stories/2004/10/warm-offices-linked-fewer-typing-errors-higher-productivity.

[85] Hirsch, J. and Associates, "eQUEST: the QUick Energy Simulation Tool," http://www. doe2. com/equest/, 2010.

[86] Ray Daniel Zimmerman, Carlos Edmundo Murillo-Sánchez, and Robert John Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, 2011.

[87] Qinglai Wei, Derong Liu, Guang Shi, and Yu Liu, "Multibattery optimal coordination control for home energy management systems via distributed iterative adaptive dynamic programming," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 7, pp. 4203–4214, 2015.

[88] Derong Liu, Qinglai Wei, Ding Wang, Xiong Yang, and Hongliang Li, *Adaptive Dynamic Programming with Applications in Optimal Control*, Springer, 2017.

[89] Glen I. Earthman, "School facility conditions and student academic achievement," Tech. Rep., UCLA's Institute for Democracy, Education, & Access, 2002.

[90] Jun Hao, Yi Gu, Jun Jason Zhang Yingchen Zhang, and David Wenzhong Gao, "Locational Marginal Pricing in the Campus Power System at the Power Distribution Level (In print)," in *Power and Energy Society General Meeting, 2016 IEEE*. IEEE, 2016, pp. 1–5.

[91] Paul J Werbos, "Approximate dynamic programming for real-time control and neural modeling," *Handbook of intelligent control*, pp. 493–526, 1992.

[92] Lei Yang, Russell Enns, Yu-Tsung Wang, and Jennie Si, "Direct neural dynamic programming," *Stability and control of dynamical systems with applications*, pp. 193–214, 2003.

[93] Xiong Yang, Derong Liu, and Qinglai Wei, "Online approximate optimal control for affine non-linear systems with unknown internal dynamics using adaptive

dynamic programming," *IET Control Theory & Applications*, vol. 8, no. 16, pp. 1676–1688, 2014.

[94] Derong Liu and Panos J Antsaklis, *Stability and Control of Dynamical Systems with Applications: A Tribute to Anthony N. Michel*, Springer Science & Business Media, 2012.

[95] Anders Rantzer, "Relaxed dynamic programming in switching systems," *IEE Proceedings-Control Theory and Applications*, vol. 153, no. 5, pp. 567–574, 2006.