Spatiotemporal graphical modeling for cyber-physical systems

by

Linjiang Wu

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Mechanical Engineering

Program of Study Committee:
Soumik Sarkar, Major Professor
Sourabh Bhattacharya
Anuj Sharma

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this thesis. The Graduate College will ensure this thesis is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2018

DEDICATION

I would like to dedicate this thesis to my major professor, Dr. Soumik Sarkar. Without his constant support, patience, and expert guidance throughout my graduate study, I would never have been able to complete this work and become more skilled in this field. I would also like to dedicate this thesis to my wife Wanxi Peng and to my son Jason Zhexi Wu for their unconditional love and support when completing this work.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my thanks to everyone who has inspired me and helped me with various aspects in completing this work. First and foremost, I would like to express my sincerest gratitude and appreciation to Dr. Soumik Sarkar for his immeasurable amount of guidance, patience and support throughout my graduate study and research. His insights and words of encouragement are truly a blessing to me and have often inspired me to become an excellent, well-rounded practitioner of science. He is certainly a great advisor, a thoughtful friend, and a role model. I would also like to thank my committee members, Dr. Sourabh Bhattacharya and Dr. Anuj Sharma, for providing insightful advice, efforts and contributions to this work. I am especially grateful to my family for constantly offering emotional support while enrolled in the graduate program. I would also like to thank my friends, especially my labmates, who have been the best friends to get help from, share the joys of successful achievements and the griefs of having papers rejected for publication. Thank you all for being there with me during both good and bad times, and for making my time as a graduate student extremely delightful and memorable. I would like to thank all of my collaborators. Without their invaluable opinions and expertise, I would never have produced tangible results from my research efforts. The work has been supported in part by the National Science Foundation under Grant No. CNS-1464279, National Science Foundation Partnerships for Innovation: Building Innovation Capacity (PFI:BIC) program under Grant No. 1632116, and Iowa DOT Office of Traffic Operations Support Grant.

ABSTRACT

Cyber-Physical Systems (CPSs) are combinations of physical processes and network computation. Modern CPSs such as smart buildings, power plants, transportation networks, and power-grids have shown tremendous potential for increased efficiency, robustness, and resilience. However, such modern CPSs encounter a large variety of physical faults and cyber anomalies, and in many cases are vulnerable to catastrophic fault propagation scenarios due to strong connectivity among their sub-systems. To address these issues, this study proposes a graphical modeling framework to monitor and predict the performance of CPSs in a scalable and robust manner.

This thesis investigates on two critical CPS applications to evaluate the effectiveness of this proposed framework, namely (i) health monitoring of highway traffic sensors and (ii) building energy consumption prediction. In highway traffic sensor networks, accurate traffic sensor data is essential for traffic operation management systems and acquisition of real-time traffic surveillance data depends heavily on the reliability of the physical systems. Therefore, detecting the health status of the sensors in a traffic sensor network is critical for the departments of transportation as well as other public and private entities, especially in the circumstances where real-time decision making is required. With the purpose of efficiently determining the traffic network status and identifying failed sensor(s), this study proposes a cost-effective spatiotemporal graphical modeling approach called spatiotemporal pattern network (STPN). Traffic speed and volume measurement sensors are used in this work to formulate and analyze the proposed sensor health monitoring system. The historical time-series data from the networked traffic sensors on the Interstate 35 (I-35) within the state of Iowa is used for validation. Based on the validation results, this study demonstrates that the proposed graphical modeling approach can: (i) extract spatiotemporal dependencies among the different sensors which lead to an efficient graphical representation of the sensor network in

the information space, and (ii) distinguish and quantify a sensor issue by leveraging the extracted spatiotemporal relationship of the candidate sensor(s) to the other sensors in the network.

In the building energy consumption prediction case, we consider the fact that energy performance of buildings is primarily affected by the heat exchange with the building outer skin and the surrounding environment. In addition, it is a common practice in building energy simulation (BES) to predict energy usage with a variable degree of accuracy. Therefore, to account for accurate building energy consumption, especially in urban environments with a lot of anthropogenic heat sources, it is necessary to consider the microclimate conditions around the building. These conditions are influenced by the immediate environment, such as surrounding buildings, hard surfaces, and trees. Moreover, deployment of sensors to monitor the microclimate information of a building can be quite challenging and therefore, not scalable. Instead of applying local weather data directly on building energy simulation (BES) tools, this work proposes a spatiotemporal pattern network (STPN) based machine learning framework to predict the microclimate information based on the local weather station, which leads to better energy consumption prediction in buildings.

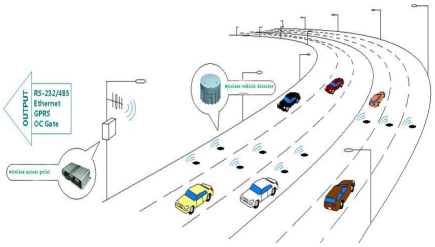
CHAPTER 1. INTRODUCTION

1.1 Motivation

The term cyber-physical systems refers to systems (such as in Fig. 1.1) with integrated physical and computational capabilities that can communicate with humans under certain modalities Baheti and Gill (2011). Over the past years, the concepts of CPS have been widely applied in medical devices and systems Lee et al. (2012), transportation systems Kuo and Szeto (2018), robotics Pasqualetti et al. (2018); Zhang et al. (2018), integrated buildings Liu et al. (2018), and other critical infrastructures to improve the system efficiency, safety, and energy savings. For example, transportation systems are able to benefit considerably from better embedded intelligence in automobiles and traffic routes, which improves safety and efficiency Qu et al. (2010). Similarly, constructed integrated building control systems (such as HVAC and lighting) significantly improve energy efficiency and supply variability Guan et al. (2010).

Due to their safety- and time-critical roles, CPSs are required to be robust and resilient. In this context, the hardware robustness and information integrity should be monitored and evaluated constantly. There are varieties of existing techniques have been applied on hardware status monitoring and evaluating, such as multiple sensor cross validation Sallans et al. (2005), periodic calibration Vico et al. (2015), and repeat experiments Wu et al. (2018a). Such methods can be easily applied on small systems like medical machines Lee et al. (2012), critical sensors applied in industry Basir and Yuan (2007), and simple repeating actions like robots on an assembly line Song et al. (2000). However, for large fundamental CPSs such as traffic systems and integrated buildings, it is too expensive, in labor and equipment, to apply the above techniques to health monitoring and energy performance prediction Wang et al. (2018). In previous research, there are some applications have been used for regional or partial traffic sensor network analysis, health monitoring, and event detection, such as macroscopic traffic modeling Thonhofer et al. (2018) and virtual reference feedback

CPS for integrated building
https://goo.gl/images/zxp2Yp



CPS for freeway traffic system
https://goo.gl/images/r4Bp2V

CPS for medical machine
https://goo.gl/images/Mz6jGp

Figure 1.1 *Examples of large distributed cyber-physical system*

tuning (VRFT) Jin et al. (2014). Some widely used machine learning techniques, such as support vector machines (SVM) Amasyali and El-Gohary (2018) and artificial neural network (ANN) González and Zamarreno (2005), have been used for energy performance prediction for some special buildings like laboratories and critical operation rooms. But for most communities and commercial buildings in the U.S, it is either expensive or technically challenging to predict building energy consumption. To resolve these concerns, this thesis proposes a spatiotemporal graphical modeling methodology (will be discussed in chapter 2) for CPSs to monitor traffic networked sensors as well as to predict the building energy consumption performance (HVAC systems).

1.2 Related Work and Contributions

As CPSs play a critical role in many engineering processes, it is not unusual for modern distributed CPSs to experience a large variety of anomalies. When overlooked, these CPSs are vulnerable to catastrophic fault propagation due to the strong connectivity between dierent sub-systems.

Monitoring the health conditions and predicting the performance of these CPSs become highly intractable due to the complex mechanism of fault propagation with a large variety of operating modes. In this section, there are two typical cyber-physical systems have been highlighted for further study, large distributed traffic sensor network and integrated building. Section 1.2.1 explains the existing techniques that have been applied on large distributed sensor network health monitoring or anomaly detection, and the approach this study proposes. In section 1.2.2, this work presents related works that focus on building energy consumption prediction, especially for HVAC systems, and the advantages of applying data-driven techniques for performance prediction which is proposed in this thesis.

1.2.1 Large Distributed Sensor Network Health Monitoring

With increasing dependencies on sensors for condition monitoring, diagnostics, and decision making in large infrastructure systems Wenjie et al. (2005); Wang et al. (2011), the reliability of the sensors themselves is critical in terms of collecting accurate information from the system of interest. Anomalous sensors may misrepresent the real conditions and mislead operators and management systems that may result in incorrect decision making, performance degradation, damage to properties, waste in energy, and even harm to the health of human beings. Therefore, there are many studies on sensor health monitoring and calibration, such as fault detection and diagnose (FDD) Wang and Chen (2004) and principle component analysis (PCA) based FDD Xiao et al. (2006) have been proposed for HVAC online sensor monitoring. Deep Belief Networks (DBN) Tamilselvan et al. (2011) have been applied to multi-sensor health diagnosis, and a number of methods have been applied for sensor accuracy calibration and validation, especially for biosensors studied in Luo et al. (1998); Brimacombe et al. (2009). However, most of the former studies tend to use sensor redundancy approaches by considering one data source as the ground truth or physical experiments to validate another data source Sallans et al. (2005). Such systems typically have multiple collocated sensors to monitor the critical points Jeong et al. (2006); Harris et al. (1995), which may be reasonable for expensive, safety-critical systems or small systems where only

limited monitoring points are needed Bhuiyan et al. (2009). In practice, such an approach may not be feasible in large distributed sensor network systems such as a large transportation network Liu et al. (2016b), which may not have multiple data sources to cross-validate the data they obtained or have limited budget for sensor implementation. Thus, these systems are too difficult for operators to take physical experiments for each sensor's accuracy calibration.

In chapter 3 (published in International Journal of Prognostics and Health Management Wu et al. (2018b)), a graphical modeling technique based on spatiotemporal pattern networks (STPN) has been proposed for traffic sensor network off-line and online health monitoring. In this application, year long real traffic data has been applied to validate this proposed framework and its results have been used to compare with a benchmark method, called traffic flow theory Wells et al. (2008); Dailey (1999), proposed based on an inherent relationship between speed, volume, and occupancy. The results show that the proposed spatiotemporal graphical modeling framework is cost-effective and robust for large distributed sensor network health monitoring. Specific contributions of chapter 3 are as follows:

a. Background and preliminaries of spatiotemporal pattern networks (STPN) based on information theoretic metric for causality and inference based metric for root cause analysis.

b. Benchmark method, online detection framework and off-line detection framework are applied to a real traffic sensor network problem based on 10 traffic sensors on Interstate 35 highway in Iowa.

c. Detailed results are presented using tables and figures, and chapter is concluded with discussions on the performance of each method..

1.2.2 Integrated Building Energy Consumption Prediction

According to the World Energy Outlook 2017 IEA. (2017), the world wide demand for energy consumption is still rapidly growing, which raises concerns about energy supply difficulties and environment changes (eg. climate change). Among various sectors, buildings have become a major consumer of energy and have a total share of 40% energy usage in the United States in 2016, as reported by the U.S. Energy Information Administration (EIA). The energy consumption in

buildings is influenced by many factors, such as ambient weather conditions, building materials and structure, and the operation of sub-systems like lighting and HVAC systems. Among these impact factors, ambient weather information is the most difficult to be obtained, which makes it difficult to predict building energy consumption accurately. Therefore, this work proposes a spatiotemporal pattern network (STPN) based graphical modeling technique to obtain the ambient weather conditions through a local weather station.

Ambient weather conditions are the main factors affecting the heat exchange from surroundings to the inside of building. Commonly, the surrounding environment information of a building is considered to be from a meteorological weather data Meek and Hatfield (1994) or a local micro weather station and is used in various building energy simulation (BES) tools for energy consumption or energy costs analysis Swan and Ugursal (2009). Given the fact that the microclimatic conditions around the building are affected by the neighboring buildings, vegetation, air flow patterns, etc. These factors also affect the energy performance of the building. The local weather station climatic data, as recorded in the far-field, may not account the local climatic information around the building. Hence, it is important to include microclimate information to have a better understanding of the building energy performance. The important microclimatic variables are average surrounding temperature, wind, humidity, solar radiation, and vegetation. These variables are usually monitored around the building. For such monitoring, a spatiotemporal relationship with the local weather data can be constructed as a model to predict the microclimatic condition.

In terms of the impact of microclimate conditions, this thesis proposes a data-driven based graphical modeling technique for microclimate conditions predictions (published in the 5th International High Performance Buildings Conference at Purdue), rather than predicting the energy consumption performance directly from local weather station data. In chapter 4, the proposed framework has been deployed and applied to a real data collected at interlock house in Ottumwa, Iowa. The specific contributions are:

a. A review of the importance of using microclimate information for energy consumption estimation associated with previously proposed techniques.

b. A formulation of data-driven methodologies for this application based on STPN and neural networks (NN).

c. The application of the proposed techniques and validation results of energy consumption prediction.

CHAPTER 2. SPATIOTEMPORAL PATTERN NETWORKS

2.1 Introduction

Spatiotemporal pattern networks (STPNs) is a causal graphical modeling concept followed by unsupervised learning of various system-level nominal patterns Fiore et al. (2013). By definition, STPNs represent temporal dynamics of each subsystem and their relational dependencies as probabilistic finite state automaton (PFSA) Sarkar et al. (2014). Patterns generating from individual subsystem and their relational dependencies are called atomic patterns (AP) and relational patterns (RP), respectively. Atomic patterns and relational patterns in the framework are extracted by applying symbolic dynamic filtering (SDF) and probabilistic finite state automaton (PFSA) techniques Ray (2004); Sarkar et al. (2014) as presented in section 2.2.1 & 2.2.2. Subsystems, APs and RPs are named as nodes, self-loop links, and links between pairs of nodes, respectively. This chapter mainly introduces the formation of STPNs and the features used for anomaly detection or event prediction.

2.2 Background on Spatiotemporal Pattern Networks (STPNs)

2.2.1 Symbolic Dynamic Filtering

Symbolic dynamic filtering (SDF) has been recently shown to be extremely effective for extracting key textures from time-series data for anomaly detection and pattern classification Rao et al. (2009). The basic concept of SDF is built based on the concept of *Symbolic Dynamics, Information Theory, and Statistical Signal Processing*, where time-series data from target system in the fast time scale are analyzed at discrete epochs in the slow time scale (as shown in Fig. 2.1). An important feature of SDF is capturing the relevant statistics by converting the time-series data into a symbol sequence using appropriate coarse-graining of the embedded information. The core idea is that a

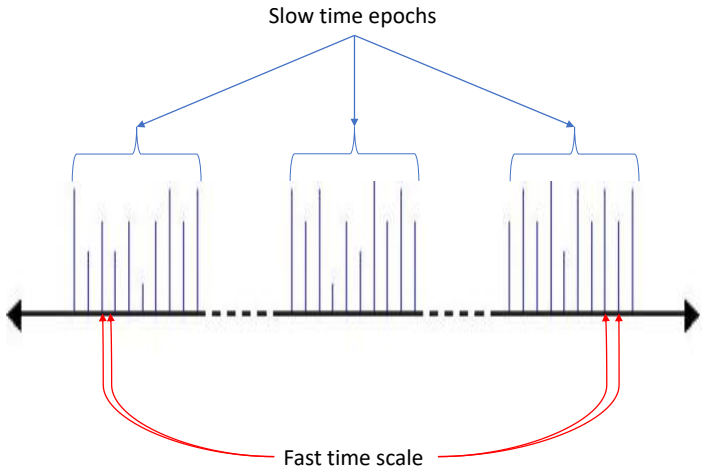


Figure 2.1 *Illustration of two-time-scale dynamics, the definition of slow time is the symbolized epoch*

symbol sequence (i.e., discretized time-series) emanated from a symbolization process, shown in Fig. 2.2 can be approximated by a Markov chain of order D (also called depth), called D-Markov machine Sarkar et al. (2014) that captures key behavior of the underlying process.

The symbolization process (also called partitioning Ray (2004); Sarkar et al. (2013a)) is as follows: Let $\mathbb{X}$ represents a set of partitioning functions, $\mathbb{X}: X(t) \rightarrow S$, which can transform a general dynamic system (time-series $X(t)$) into a symbol sequence S using an alphabet set Σ. Researchers have proposed different partitioning approaches according to different objective functions, such as uniform partitioning (UP), maximum entropy partitioning (MEP), statistically similar discretization (SSD) Sarkar and Srivastav (2016), and maximally bijective discretization (MBD) Sarkar et al. (2013b). The detailed introduction of statistical similarity-based discretization (SSD) and maximally bijective discretization (MBD) algorithms are presented in chapter 3 and in Sarkar et al. (2013b).

2.2.2 Probabilistic Finite State Automaton

Probabilistic finite state automaton (PFSA) is defined in the symbolic space as presented in section 2.2.1 that the time-series data is discretized into symbol sequences and then state sequences, which is the essential part in STPNs for patterns extraction and transition matrix generation, and

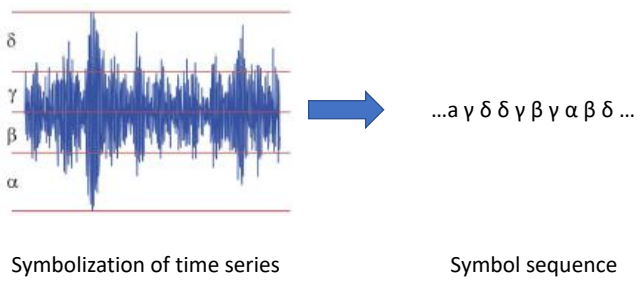


Figure 2.2 *Symbolization of measured time-series (discretization)*

it generalizes the concept of Markov chain. The D-Markov machine, as mentioned in section 2.2.1, is represented by a probabilistic finite state automaton (PFSA) that is constituted by states (represents different parts of the data space) and probabilistic transitions among these states that can be captured from time-series data. Theoretically, PFSA is constructed based on a deterministic finite state automaton (DFSA) $\mathbb{D} = (\Sigma, Q, \hat{\delta})$ Sarkar et al. (2014) as a pair $\mathbb{P} = (\mathbb{D}, \phi)$, i.e., the PFSA is a 4-tuple $\mathbb{P} = (\mathbb{D}, \phi)$, where:

1. Q is a non-empty finite state set;
2. Σ is a non-empty finite alphabet set;
3. $\hat{\delta}$: $Q \times \Sigma \to Q$ is the mapping for state transition;
4. ϕ: $Q \times \Sigma \to [0, 1]$ is the symbol generation function, i.e., probability morph function which satisfies the condition $\sum_{\sigma \in \Sigma} \phi(q, \sigma) = 1, \forall q \in Q$, and p_{ij} denotes the probability of the symbol $\sigma_j \in \Sigma$ occurring with the state $q_i \in Q$.

The framework of PFSA is shown in Fig. 2.3 and the detailed definitions of PFSA, D-Markov machine, xD-Markov machine and the learning schemes can be found in Sarkar et al. (2014); Mukherjee and Ray (2014).

Based on the above setup, the spatiotemporal pattern networks (STPNs) is formulated as below Liu et al. (2016a).

Definition. A PFSA based STPN is a 4-tuple defined as $W_D \equiv (Q^a, \Sigma^b, \Pi^{ab}, \Lambda^{ab})$, (a, b are nodes of the STPN)

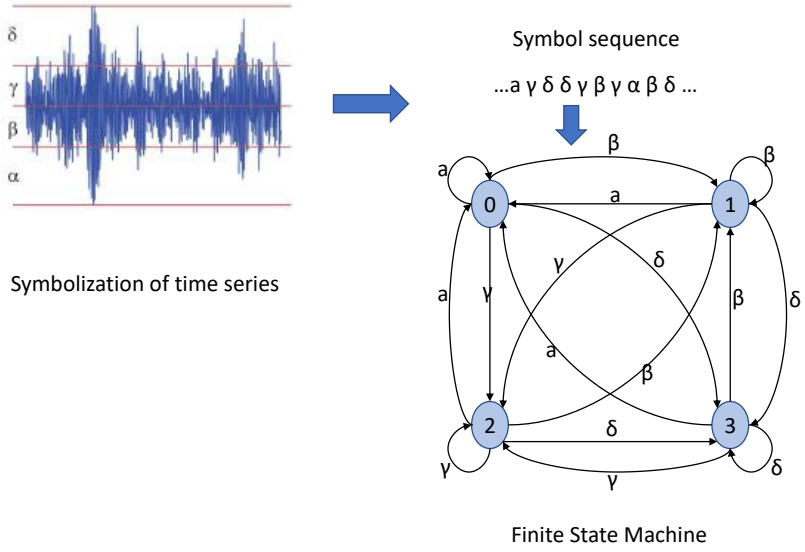


Figure 2.3 *Formation of finite state automaton for symbolic dynamic filtering*

1. $Q^a = \{q_1^a, q_2^a, \cdots, q_{|Q^a|}^a\}$ is the state set corresponding to symbol sequences S^a;

2. $\Sigma^b = \{\sigma_0, \cdots, \sigma_{|\Sigma^b|-1}\}$ is the alphabet set of symbol sequence S^b;

3. Π^{ab} is a $|Q^a| \times |\Sigma^b|$ symbol generation matrix, the ij^{th} element of Π^{ab} represents the probability of observing the symbol σ_j in the symbol list S^b while making a transition from the state q_i^a in the symbol sequence S^a; self-symbol generation matrices are called atomic patterns (APs) i.e., when $a = b$, cross-symbol generation matrices are called relational patterns (RPs) i.e., when $a \neq b$.

4. Λ^{ab} is a metric that can represent the importance of the learnt pattern (or degree of causality) for $a \rightarrow b$ which is a function of Π^{ab}.

An illustration of STPN is shown in Fig. 2.4.

2.3 Information Theoretic Metric for Causality Using STPNs

Based on the above definition of STPN, we can use the atomic or relational patterns to interpret the causal dependencies among the sensors. In this context, information theoretic criteria have

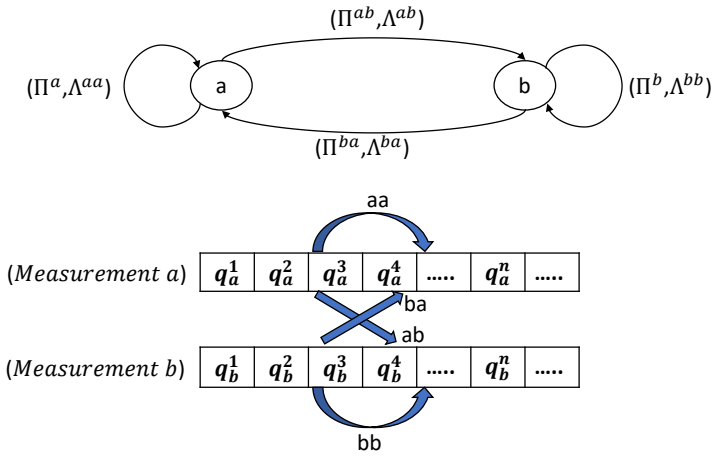


Figure 2.4 *Demonstration of Spatiotemporal Pattern Network (STPN) for extraction of atomic and relational patterns (using D-Markov and xD-Markov machines respectively and $D = 1$, i.e., states and symbols are equivalent).*

been widely used, e.g., transfer entropy Wibral et al. (2011) and mutual information Sarkar et al. (2014); Solo (2008). In this study, the concept of mutual information Λ is applied for representing the importance of the patterns (APs & RPs). The definition of Λ^{aa} and Λ^{ab} are as follows.

$$\Lambda^{aa} \triangleq I^{aa} = I(q^a_{k+1}; q^a_k) = H(q^a_{k+1}) - H(q^a_{k+1}|q^a_k) \tag{2.1}$$

where, I^{aa} is the mutual information of atomic pattern (a, a), H is the conditional entropy defined as follows,

$$H(q^a_{k+1}) = -\sum_{i=1}^{Q^a} P(q^a_{k+1} = i) \log_2 P(q^a_{k+1} = i)$$

$$H(q^a_{k+1}|q^a_k) = \sum_{i=1}^{Q^a} P(q^a_k = i) H(q^a_{k+1}|q^a_k = i)$$

$$H(q^a_{k+1}|q^a_k = i) = -\sum_{j=1}^{Q^a} P(q^a_{k+1} = j|q^a_k = i)$$

$$\cdot \log_2 P(q^a_{k+1} = j|q^a_k = i)$$

Here, I^{aa} essentially captures the temporal self-prediction capability of the sensor node a and q^a_k denotes the k^{th} state of node a. Similarly, the mutual information for the relational pattern (a,b)

can be expressed as:

$$\Lambda^{ab} \triangleq I^{ab} = I(q^b_{k+1}; q^a_k) = H(q^b_{k+1}) - H(q^b_{k+1}|q^a_k) \tag{2.2}$$

where, I^{ab} is the mutual information of pattern (a, b), H is the conditional entropy defined as follows,

$$H(q^b_{k+1}|q^a_k) = \sum_{i=1}^{Q^a} P(q^a_k = i) H(q^b_{k+1}|q^a_k = i)$$

$$H(q^b_{k+1}|q^a_k = i) = -\sum_{j=1}^{Q^b} P(q^b_{k+1} = j|q^a_k = i)$$

$$\cdot \log_2 P(q^b_{k+1} = j|q^a_k = i)$$

Detailed description of mutual information theoretic causality metric in the context of APs and RPs can be found in Sarkar et al. (2014).

2.4 Inference Based Metric Using STPNs

The mutual information theoretic metric introduced in section 2.3, requires significant amount of data for estimating the state transition probabilities and hence, may not be ideally suited for online real time decision making. Therefore, an alternative inference based metric is presented here which utilizes a short time window of data to compute the metric using a Dirchlet prior on the state trasition probabilities. To compute this metric, a two-step process is needed that includes a modeling and an inference phase Liu et al. (2016a).

In the modeling phase, the entire time-series in the nominal condition is considered, where the multivariate time-series is denoted by $X = \{X^{\mathcal{N}}(t), t \in \mathbb{N}, \mathcal{N} = 1, 2, \cdots\}$, where $\mathcal{N}$ is the number of measurement sensors in the network. The multivariate time-series is symbolized into $S = \{S^{\mathcal{N}}\}$ and then state sequences are generated with the STPN formulation, noted by $Q = \{Q^{\mathcal{N}}, \mathcal{N} = 1, 2, \cdots\}$.

In the inference phase, a short time-series is analyzed, $\tilde{X} = \{\tilde{X}^{\mathcal{N}}(t), t \in \mathbb{N}^*, \mathcal{N} = 1, 2, \cdots\}$, where $\mathbb{N}^*$ is a subset of $\mathbb{N}$. The length of the short time-series depends on the selection of a window size, which is flexible and can be overlapping. The symbolic subsequences for the short time-series is noted as $\tilde{S} = \{\tilde{S}^{\mathcal{N}}\}$, and the state sequences is noted as $\tilde{Q}$. An importance metric Λ^{ab} is defined for

a given short subsequence (described by short state subsequence $\tilde{Q}$ and short symbol subsequence $\tilde{S}$). The value of this metric suggests the importance of the pattern Π^{ab} or the degree of causality in $a \rightarrow b$ as evidenced by the short subsequence. In this context, we consider

$$\Lambda^{ab}(\tilde{Q}, \tilde{S}) \propto Pr(\{\tilde{Q}^a, \tilde{S}^b\}|\Pi^{ab}) \tag{2.3}$$

where $Pr(\{\tilde{Q}^a, \tilde{S}^b\}|\Pi^{ab})$ is the conditional probability of the joint state-symbol subsequence given the pattern Π^{ab}.

With this definition of Λ^{ab} and with proper normalization, the inference based metric $\Lambda^{ab}(\tilde{Q}, \tilde{S})$ can be obtained as follows,

$$\begin{aligned} &\Lambda^{ab}(\tilde{Q}, \tilde{S}) = \\ &K \prod_{m=1}^{|Q^a|} \frac{(\tilde{N}_m^a)!(N_m^a + |\Sigma^b| - 1)!}{(\tilde{N}_m^a + N_m^a + |\Sigma^b| - 1)!} \prod_{n=1}^{|\Sigma^b|} \frac{(\tilde{N}_{mn}^{ab} + N_{mn}^{ab})!}{(\tilde{N}_{mn}^{ab})!(N_{mn}^{ab})!} \end{aligned} \tag{2.4}$$

where, K is a proportional constant, $N_{mn}^{ab} \triangleq |\{(Q^a(k), S^b(k+1)) : S^b(k+1) = \sigma_n^b \mid \mathcal{Q}^a(k) = q_m^a\}|$, $N_m^{ab} = \sum_{n=1}^{|\Sigma^b|}(N_{mn}^{ab})$, $\tilde{N}_{mn}^{ab}$ and $\tilde{N}_m^{ab}$ are similar to N_{mn}^{ab} and N_m^{ab}, $|Q^a|$ is number of states in state sequence $\tilde{Q}^a$, and $|\Sigma^b|$ is number of symbols in symbol sequence $\tilde{S}^b$.

A detailed derivation can be found in Liu et al. (2016a). Thus, with Eq. 2.4, the inference metrics Λ^{ab} of APs (i.e., when $a = b$) and RPs (i.e., when $a \neq b$) are obtained with respect to the *short subsequences*.

2.5 Summary

In conclusion, the proposed STPN is capable of recognizing patterns of anomalous behavior by learning a model for the nominal behavior and identifying departures from such behavior due to parametric or non-parametric changes in the underlying system.

CHAPTER 3. TRAFFIC SENSOR HEALTH MONITORING USING SPATIOTEMPORAL GRAPHICAL MODELING

3.1 Introduction

Accurate traffic sensor data is essential for traffic operation management systems and acquisition of real-time traffic surveillance data depends heavily on the reliability of the traffic sensors (e.g., wide range detector, automatic traffic recorder). In a typical road transportation network, traffic sensors are deployed on freeways primarily to collect real-time data for traffic adaptive signal control and mitigating recurring or nonrecurring congestion Klein et al. (2006). According to the U.S. Department of Transportation (the U.S. DOT), the sensors are typically installed about every 2 miles and facilitating sensor redundancy is not feasible due to the sheer length of roadways that requires monitoring in each state and the cost of sensors deployed (e.g., microwave radar sensor that covers multiple lanes costs at least $6200 without the installation fee based on the costs database of the U.S. DOT in 2002 Klein et al. (2006)). Therefore, a robust and feasible approach to monitor the health status of the traffic sensors is required that does not rely on redundancy of collocated sensors.

Among the existing techniques that are used to monitor the health status of the traffic sensors, a method based on the traffic flow theory has been widely adopted Wells et al. (2008); Dailey (1999) to identify the erroneous data and anomalous sensors. In this approach, the average effective vehicle length (AEVL) is computed by defining a function $F(V, C, O)$, where V,C and O denote the traffic speed, traffic volume, and the sensor occupancy respectively. An assessment criteria can be formulated to report the error rate of the sensor by evaluating whether the AEVL value meets the criteria or not FHW (2016). Although it is fast and can be used for online monitoring, the method is built upon single lane road assumption and the accuracy is seriously affected when

applied to roads with multiple lanes (e.g., 2 lanes when only 1 lane has vehicle passing at the recording time).

In this work, we consider the fact that the traffic sensors in the same freeway and direction form a sensor network in the information space. Therefore, under nominal conditions, the data collected by these sensors should follow a stable spatiotemporal relationship among themselves that can be captured by an efficient learning technique using historical data. Such a stable relationship will be affected when one or more sensors degrade in performance. Hence, discovering the relationships among the sensors during operation with respect to the (historical) nominal conditions can provide us indications whether a sensor is healthy or not. In this context, this work applies a recently proposed spatiotemporal graphical modeling approach, called the spatiotemporal pattern network (STPN, built on the concepts of symbolic dynamics filtering, SDF) Sarkar et al. (2014); Liu et al. (2017); Jiang and Sarkar (2015), to build a novel sensor health monitoring framework for traffic sensors.

Contributions: The main technical contributions of this chapter are: (i) STPN learning framework to capture the causal dependencies between the sensors in a traffic sensor network using multi-model information sources such as speed, volume, and the sensor occupancy, (ii) an off-line traffic sensor anomaly detection and isolation framework using the learnt STPN model, and (iii) an on-line traffic sensor anomaly detection and isolation approach via a recently proposed STPN+ RBM (RBM: Restricted Boltzmann Machine) technique Liu et al. (2016a) is used for the first time in the present context. The proposed approaches are validated based on a real traffic sensor network, installed on Interstate 35 from Ankeny to Ames in the state of Iowa. The data set was collected by Wavetronix LLC. in Oct., Nov. and Dec. 2016, and the ground truth (nominal and anomalous categorization) has been established manually based on careful inspection and collected field images. Performance evaluation and comparison with the benchmark method (AEVL) are presented using both real traffic sensor anomalies and simulated traffic sensor degradations (for general sensor degradation or fault types Najafi et al. (2004)) in this real life setting.

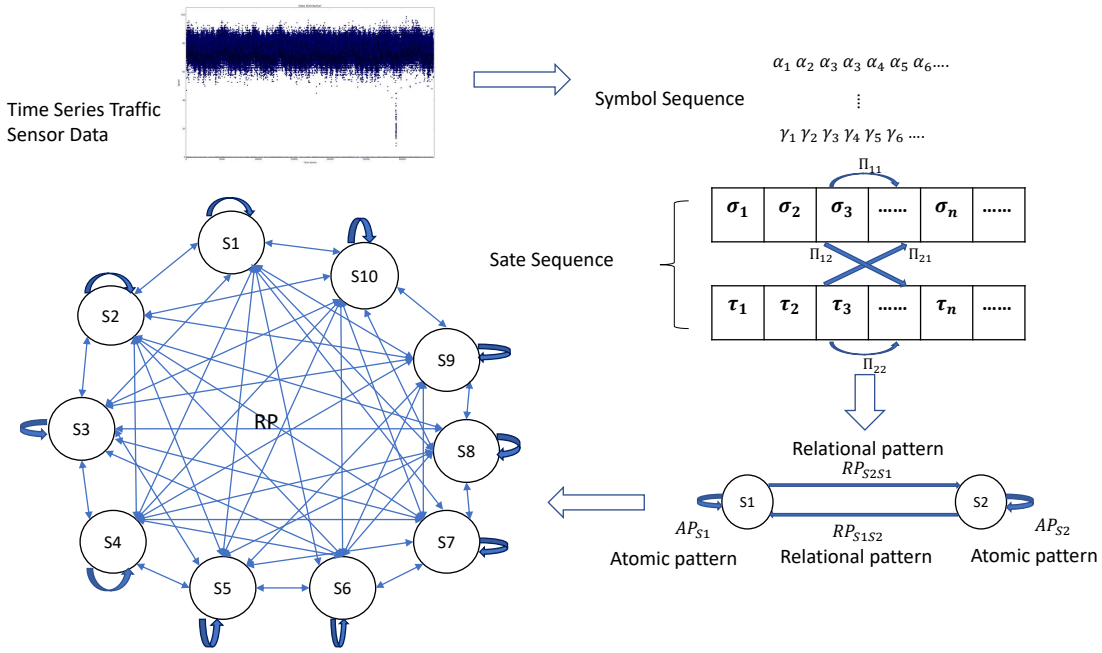


Figure 3.1 *Extraction of atomic and relational patterns (using D-Markov and xD-Markov machines respectively and $D = 1$, i.e., states and symbols are equivalent) in sensor network $\boldsymbol{S}$, where $\mathcal{N} = (1, 2, .., 10)$ sensors in the network to characterize individual sub-system behavior and interaction behavior among different sub-systems.*

Including the introduction, this chapter is presented in 5 sections. Section 2 presents the graphical modeling background, an information theoretic metric and an inference based metric leveraged to build the sensor health monitoring framework. Section 3 describes the three approaches for sensor health monitoring: (i) the benchmark method (AEVL), (ii) the proposed off-line method, using the information theoretic metric, and (iii) the proposed on-line approach, using the inference based metric. The results obtained with the three different methods are described in Section 4 and the study is summarized in Section 5.

3.2 Background on STPNs Based Graphical Modeling

Based on the framework of Spatiotemporal Pattern Networks (STPNs) described in Chapter 2, the STPNs of the above traffic sensor network can be constructed as shown in Fig. 3.1.

3.2.1 STPN Based Information Theoretic Metric

The information theoretic metric, as presented in Chapter 2, is applied based on mutual information criteria to identify the importance of an AP or an RP of the traffic sensor network. Mutual information (I) quantifies the importance of APs and RPs (e.g.,I^{aa}, I^{ab}) and is able to assign weights on the patterns. The detailed formulations are provided in Chapter 2 and in Butte and Kohane (1999).

3.2.2 Online Anomaly Detection with STPN+RBM Framework

The inference metrics (Λ^{ab}) shown in the chapter 2 can be further normalized and converted into binary states (0 for low values and 1 for high values) for APs and RPs. With multiple short subsequences, a large number of examples can be formed, whose characteristics represent the systematic behavior. Then, Restricted Boltzmann Machine (RBM) is trained to capture the most likely system-wide behavior, and detect any anomaly via identifying a low probability event.

For RBM, weights and biases are learnt so that the feature configurations observed during nominal operation of the system obtain low energy (or high probability). Consider a system state that is described by a set of visible variables $\mathbf{v} = (v_{1,2}, v_3, \cdots, v_D)$ and a set of hidden (latent) variables $\mathbf{h} = (h_1, h_2, h_3, \cdots, h_M)$. Here, the normalized inference metrics (Λ^{ab} are used as the inputs for the RBM, thus, $|\mathbf{v}| = |\Lambda^{ab}|$, i.e., the number of visible units equal to the number of patterns learned by the STPN. The variables can be binary or real-valued depending on the need. Now, each joint configuration of these variables determines a particular state of the system and an energy value $E(\mathbf{v}, \mathbf{h})$ is associated with it. The energy values are functions of the weights of the links between the variables (for RBM, internal links within the visible variables and the hidden variables are not considered) and bias terms related to the variables.

With this setup, the probability of a state $P(\mathbf{v}, \mathbf{h})$ depends only on the energy of the configuration ($\mathbf{v}$, $\mathbf{h}$) and follows the Boltzmann distribution

$$P(\mathbf{v}, \mathbf{h}) = \frac{e^{-E(\mathbf{v},\mathbf{h})}}{\sum_{\mathbf{v},\mathbf{h}} e^{-E(\mathbf{v},\mathbf{h})}} \tag{3.1}$$

$$E(\mathbf{v}, \mathbf{h}) = -\mathbf{h}^T W \mathbf{t} - B^T \mathbf{v} - \mathbf{c}^T \mathbf{h} \tag{3.2}$$

Where W is the weight of the hidden unit, B and $\mathbf{c}$ are the biases of the visible units and hidden units, respectively.

Anomaly detection process. During training, weights and biases are obtained via maximizing likelihood of the training data. During testing, short testing subsequences are converted into an $\mathbb{N}^2$-dimensional binary vectors using the same *inference phase* of the training process. Multiple testing (possibly overlapping) subsequences are applied to compute a distribution of free energy. For the nominal condition, the distribution of free energy should be close to that of the training data, while the anomalous data should differ from the nominal condition. Further details of the STPN+RBM framework can be found in Liu et al. (2016a).

3.3 Problem Setup and Methodology

3.3.1 Problem Setup

Consider a sensor network $\mathbf{S} = \{\mathcal{S}1, \mathcal{S}2, ..., \mathcal{S}\mathcal{N}\}$ with $\mathcal{N}$ sensors (as illustrated in Fig. 3.1). Each sensor is represented by the measurements in our case, which can be univariate time-series (speed) or multivariate time-series (speed, volume, and occupancy). The sensor health monitoring task is to find out the anomalous sensor(s) based on a certain performance metric $\mathcal{M}$. Therefore, the sensor health monitoring problem can be formulated as:

$$\text{Finding } \mathbf{S}^{ano} \subset \mathbf{S} \tag{3.3}$$

where $\mathbf{S}^{ano}$ is the subset of the sensor network that are anomalous.

Three performance metrics are illustrated in the following sections where the first one is based on the traffic flow theory and used as the benchmark, and the later two are proposed in this study for the purposes of off-line and on-line detection. Note that in this work, sensor health monitoring with univariate time-series (speed) is noted as the 1D model, and the one with two time-series (speed, volume) is noted as the 2D model.

3.3.2 Benchmark Method Based on Traffic Flow Theory

According to the traffic flow theory, there is an inherent relationship between speed, volume, and occupancy Dailey (1999) and hence, such relationship can be applied to assess sensor data quality. Authors in Wells et al. (2008) proposed a method to identify sensor errors via evaluating the relationship among speed, volume, and sensor occupancy. In this framework, the sum of average effective vehicle length (AEVL) and the detection range (DTR) is estimated by the following empirical rule:

$$AEVL + DTR = \frac{5280 * Speed * Occupancy}{Volume}, \tag{3.4}$$

where AEVL is in feet, DTR of the sensor is in feet, speed is in miles per hour, occupancy is a fractional number between 0 and 1 representing percentage of time when the sensor is occupied, volume is in vehicles per hour and the scalar 5280 is used for unit standardization.

We use this relationship based on our radar sensor as the benchmark method for sensor health monitoring. Note, DTR is the length of a loop detector Chen et al. (2001), a typical sensor used in this type of applications. As the data from Wavetronix HD sensors uses a virtual line to represent the detectorWav (2016), DTR equals to 0. According to FHW (2016); Minge et al. (2012), the possible distance between vehicles should fall within the range of 10 ft to 75 ft, which provides a method to monitor the health of the sensor by using the AEVL equation. A sensor is identified as erroneous when the output falls outside this range and the ratio of error counts to the total number of data points is called the error rate E_r.

3.3.3 Off-line Sensor Fault Detection Using STPN

Among the proposed STPN-based sensor monitoring solutions, we first present the off-line method that uses the information theoretic metric defined in Section 2.3. We begin the discussion with the symbolization procedures which is critical for the success of the proposed schemes.

3.3.3.1 1D data and 2D data symbolization

In 1D STPN model, we use traffic speed as the univariate input. The speed distribution from historical data shown in Fig. 3.2 demonstrates that the speed mainly falls in the range of 65 miles per hour to 83 miles per hour and is normally distributed. In this case, we found that compared with UP and MEP discretization, SSD discretization that aims to preserve the nature of the continuous data distribution in the discrete domain, is much more effective. The SSD discretization for traffic speed is based on the modified Douglas-Peucker (DP) algorithm Douglas and Peucker (2011). The detailed formulation of the SSD scheme can be found in Sarkar and Srivastav (2016), we provide a brief description here in Algorithm 1 for completeness.

Let $X_V(t) \in \Omega^1$ be a one dimensional time-series data, where interval Ω^1 is a compact subset of $\mathbb{R}$. Let $\mathbf{B}=\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, ...\}$, $\mathbf{b}_i \in \Omega^1, \mathbf{b}_i < \mathbf{b}_j$ for $i < j$ be an ordered set with discrete levels that have to be determined to discretize Ω^1. There exists a function $F(X_V(t))$ which represents the true underlying cumulative density function (CDF) of $X_V(t)$. In the discrete domain, we define the density $\tilde{F}_{\mathbf{B}}(X_V(t))$ that aims to preserve the statistical properties of X(t) in the sense of minimizing the distance d with respect to $|\tilde{F} - F|$, where $\tilde{F}$ is a piece-wise linear and continuous function. In this study, we use the Kolmogorov-Smirnov statistic test to compute the distance d, where $d = \sup_{X_V(t)} |\tilde{F}_{\mathbf{B}} - F|$. Note, generally, the true density $F(X_V(t))$ is not available, here we propose to use the empirical density $\tilde{F}_n(X_V(t))$ (observated), where n is the number of data points. With the SSD technique, partitioning of the speed data is implemented and the symbol sequences for the speed sensors in the network are generated.

Algorithm 1 Statistical Similarity-based Discretization (SSD)

Input variable X_V

Input tolerance ξ on d for univariate discretization of X_V.

Compute empirical $\tilde{F}_n(X_V)$ of X_V as the set $\{(X_V^k, p^k)|Prob(X_V < X_V^k) = p^k, k = 0, 1, 2, 3, ..., K\}$ *and* $p^K = 1$

fit a line segment L_1 through the endpoints (X_V^0, p^0) and (X_V^K, p^K)

Initialization of the line segments $\mathbb{L} = \{L_1\}$

Find a split point k_s to maximize the distance between L_1 and cumulative density function $\tilde{F}_n(X_V)$, k_s is the split point.

while d$\geq \xi$ **do**

    for `i in k` **do**

        Generate two new line segments L_l and L_r corresponding to the split point k_s for L_i

        Update $\mathbb{L} = \mathbb{L} \cup \{L_l, L_r\} \backslash L_i$

    end for

    Find the k_s for segments $\mathbb{L}$ and $\tilde{F}_n(X_V)$ have the maximum distance, where the bin boundary of **B** is determined.

end while

In this work, we also construct STPN model for 2D data, where traffic speed and volume are used (data are in 20s and from the same data source as the 1 dimensional model), the correlation between speed and volume can be viewed in the figure presented in Fig. 3.3, The distribution of volume is shown in Fig. 3.4. To implement the joint (2D) symbolization of the correlated variables speed and volume, we adopt a two-step technique, SSD followed by MBD. The main idea of MBD is as in Algorithm 2.

Let $X_C(t)$ denote the volume and $X_V(t)$ denote the speed and let $\mathbb{L}$=(l_1,l_2,l_3,...) denote the set of discrete bins for volume $X_C(t)$ after implementing the SSD discretization. The main objective of the MBD scheme is to find the bins for the speed variable $X_V(t)$ which is denoted by $\mathbf{B}$=$\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, ...\}$,

such that there is a maximum possible one-to-one correspondence (hence, maximally bijective) between the bins for $X_V(t)$ and those for $X_C(t)$. In this work, the traffic volume $X_C(t)$ which has been symbolized with SSD technique is the input variable, the basis of MBD discretization as mentioned in chapter 2.

In this context, we call bin l_i corresponds to bin $\mathbf{b}_j$, i.e., $l_i \Rightarrow \mathbf{b}_j$ if $i = arg\max_h P(l_h|\mathbf{b}_j \in \mathbf{B})$, $h \in (1, .., |\mathbb{L}|)$, where $P(\cdot)$ is a probability function, l_i is the i^{th} bin for volume X_C and $\mathbf{b}_j$ is the j^{th} bin for speed X_V. Then, a reward function is defined for the discretization as follows: $R(\mathbf{b}_j) = P(l_i|\beta \in X_V)$ s.t. $\mathbb{L} \Rightarrow \mathbf{B}$. A higher reward value means that an existing (SSD) bin better corresponds to an MBD bin. The total expected reward value can be calculated by: $TR(\mathbf{B}) = \int_\beta R(\mathbf{b}_j)P(\beta)d_\beta$. In the MBD scheme, the goal is to maximize the total reward TR.

Algorithm 2 Maximally Bijective Discretization(MBD)

1: Speed X_V is the output
2: β=min(X_V)
3: k=1, initial index value
4: **while** $r_j < \max X_V$ **do**
5: Select l_i such that $P(l_i|\beta) = P_{max}(\beta)$
6: Select l_j such that $P(l_j|\beta + d_\beta) = P_{max}(\beta + d_\beta)$
7: l_i,l_j represent the existing i^{th},j^{th} bins for volume X_C
8: **if** $i \neq j$ **then**
9: $\mathbf{b}_k = \beta$, $\mathbf{b}_k$ denotes the k^{th} bin boundary of X_V.
10: k← k+1
11: **end if**
12: $\beta \leftarrow \beta+\beta$
13: **end while**
14: bin boundaries for MBD, $\mathbf{B}$=$\{\mathbf{b}_1, \mathbf{b}_2, ...\mathbf{b}_{k-1}\}$ is obtained

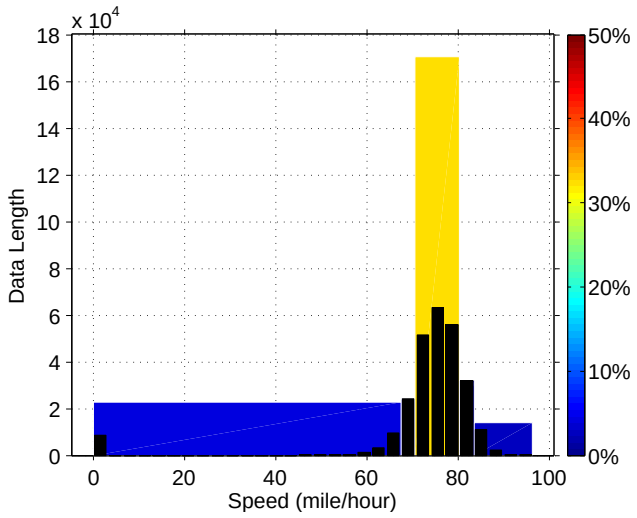


Figure 3.2 *Speed distribution in histogram which shows that the speed of 10 sensors is mainly in range from 65 mile/hour to 83 mile/hour, traffic congestion or no vehicle passing leads to 0 mile/hour.*

Algorithm 2 aims to achieve the maximally bijective discretization of $X_V(t)$, where $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, ...\}$ denotes the set of MBD bin boundaries and σ_j^V is the symbol with boundaries $\mathbf{b}_j$ and $\mathbf{b}_{j+1}$:

3.3.3.2 Fault Detection

After symbolization, multiple sets of symbol sequences $S^1, S^2, ..., S^{10}$ (representing training data and testing data respectively, and corresponding to sensor network shown in Fig. 3.1) are generated in time-series based on the training symbols Σ. Thus, based on section 2.3, the mutual information matrices (10 by 10 matrix) of the training data and testing data in the network can be obtained as:

$$\underline{\Lambda}_{trg} = (I^{S1S1}, I^{S1S2}, ..., I^{S10S10})$$

$$\underline{\Lambda}_{tst} = (I^{S1S1}, I^{S1S2}, ..., I^{S10S10})$$

Where $\underline{\Lambda}_{trg}$ and $\underline{\Lambda}_{tst}$ are the training and testing mutual information matrices respectively.

The difference $\Delta\underline{\Lambda}$ between $\underline{\Lambda}_{trg}$ and $\underline{\Lambda}_{tst}$, can be used to detect and isolate the anomalous sensor(s) in the sensor network.

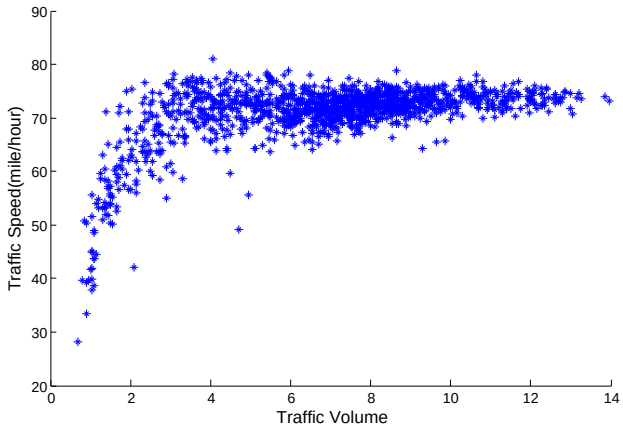


Figure 3.3 *Relationship between speed and volume, it reasonably explains the high correlations between speed and volume.*

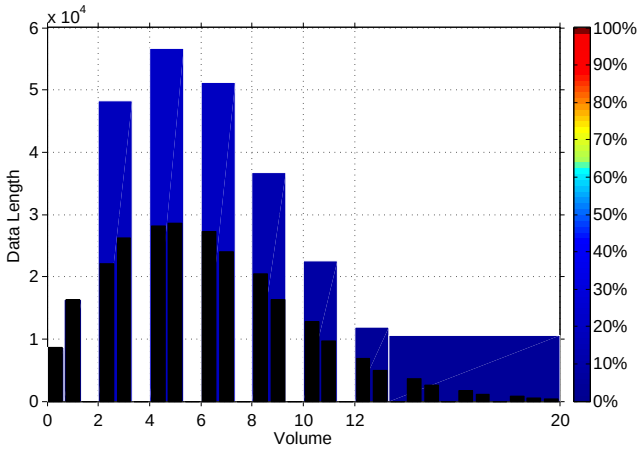


Figure 3.4 *Volume distribution in histogram. In 20s timestamp data, the volume is mainly distributed in range from 0 to 12.*

3.3.4 Online Detection with Inference Based on Spatiotemporal Graphical Modeling

With STPN+RBM framework presented in Section 2.4, an anomaly is detected as a high energy (low probability) event, and the distribution of free energy in anomalous condition differs from that in the nominal condition. Then a sequential state switching method can be used to further localize the fault in the sensor network. The idea for sequential state switching is to find potential pattern(s) that, if changed, can transition the system from a high to a low energy state. The probabilities of AP and RP's existence are discovered by the STPN, and an anomaly will influence the causality of specific patterns. Hence, by switching/flipping a pattern, its contribution on the energy status of the system can be estimated and attributed to a possible fault.

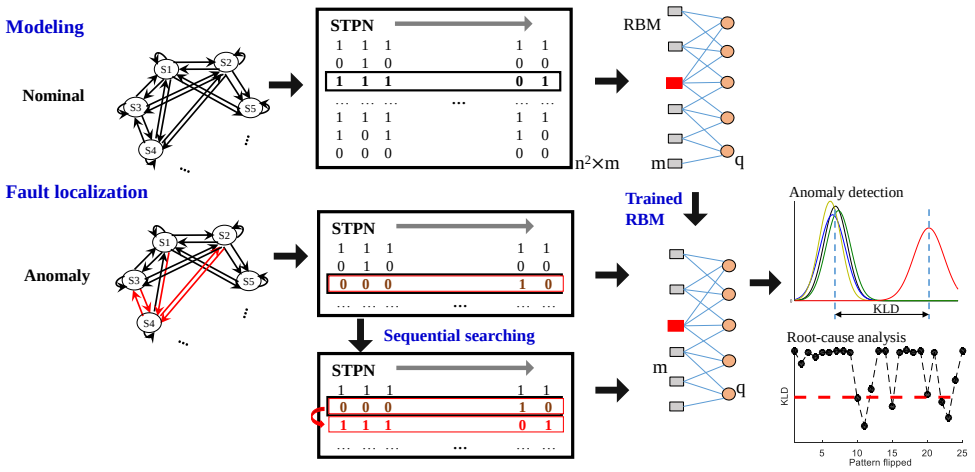


Figure 3.5 *Online detection via sequential state switching approach based on STPN+RBM framework. Spatiotemporal features are extracted from both nominal and anomalous data, with online STPN inference. Multiple sub-sequences of APs and RPs form input vectors to the RBM. Here, the RBM is only trained with nominal data, and the anomalous data is used as input to compute free energy. Anomaly is detected by identifying its high energy state. In the root-cause analysis phase (fault detection/isolation), the potentially failed patterns are obtained via evaluating free energy of the system with the perturbations.*

With the weights and biases of RBM using training data as presented in chapter 2, free energy can be computed Hinton (2012):

$$G(v) = -\sum_i v_i a_i - \sum_j \ln(1 + e^{b_j + \sum_i v_i w_{ij}})$$

The free energy in nominal conditions is noted as $\tilde{G}$. In fault conditions, a failed pattern will shift the energy from a lower state to a higher state. Assume that the patterns can be categorized into two sets, $\mathbf{v}^{nom}$ and $\mathbf{v}^{ano}$. By flipping the set of anomalous patterns $\mathbf{v}^{ano}$, a new expression for free energy is obtained:

$$\begin{aligned} G^s(v) = &-\sum_g v_g a_g - \sum_j \ln(1 + e^{b_j + \sum_g v_g w_{gj}}) \\ &-\sum_h v_h^\star a_h - \sum_j \ln(1 + e^{b_j + \sum_h v_h^\star w_{hj}}), \\ &\{v_g\} \in v^{nom}, \{v_h^\star\} \in v^{\star,ano} \end{aligned} \tag{3.5}$$

Here, $v^\star$ has the opposite state to v and represents that the probability of the pattern has been significantly changed. In this work, the probabilities of the patterns are binary (i.e. 0 or 1). Hence, we have that $v^\star = 1 - v$. The sequential state switching is formulated by finding a set of patterns v^{ano} via $\min(G^s(v^{ano}, v^{nom}) - \tilde{G})$. Flowchart of sequential state switching method is shown in Fig. 3.5.

Note that the sequential state switching method is pattern based, and the anomalous patterns are associated to the candidate nodes (sensors in this case) using a sequential search method, to find a subset $\hat{\mathbf{X}}$ of $X = \{X^{\mathcal{N}}(t),\ t \in \mathbb{N},\ \mathcal{N} = 1, 2, \cdots, 10\}$ that can interpret all of the anomalous patterns $\mathbf{\Lambda}^{ano}$. For example, a pattern $N_i \rightarrow N_j$ is identified as failed, and it indicates that the two nodes (N_i, N_j) are potentially failed. If multiple patterns from or to a node are detected as anomalous, the node is more reliable to be classified as anomalous. Thus, the node inference can be carried out via computing the anomaly score of each node. Also, as the failed patterns contribute differently to the system (in terms of energy increase in RBM, weights of failed patterns are defined. The weights of failed patterns can be formed by the difference of free energy with and without the failed pattern and then can be associated to the anomaly scores for the patterns.

With sequential state switching method, the fault sensor(s) $\mathbf{S}^{ano}$ in the sensor network can be identified and the performance metric $\mathcal{M}$ is formed using the anomaly score(s) AS, which is a measure of the energy increase (of the candidate sensor) in the RBM. By defining a threshold AS^{thres}, the on-line detection approach for the sensor network is formulated as:

$$\text{Finding a subset } \mathbf{S}^{ano} \subset \mathbf{S}, \text{ where } AS > AS^{thres} \tag{3.6}$$

Note that, as the sequential state switching method is built upon short sequences, it only needs short time-series and is suitable for online detection.

3.4 Results and Discussions

While we collected real data from Iowa interstate traffic scenarios (a network of 10 sensors on Interstate 35 from Ankeny to Ames in the state of Iowa, the data set was collected by Wavetronix LLC. in Oct., Nov. and Dec. 2016) to validate our proposed technique, actual sensor faults are somewhat rare. Therefore, it becomes difficult to use only that data for comprehensive validation. However, representative sensor degradations of different types and severity levels can be artificially injected in real nominal data and that is how we begin presenting our results.

3.4.1 Simulation Results

Based on the collected real data, the sensor faults are simulated in two ways: (i) drift–the measured speeds of the sensor(s) are artificially modified by adding different levels of bias and (ii) noise–the measurements of the sensor(s) are contaminated by a predefined level (variance) of Gaussian noise.

The original data is initially divided into a training, $X_{trg}(t)$ and testing, $X_{tst}(t)$ data for validating with real data. However, initially the training set $X_{trg}(t)$ is further divided into $\hat{X}_{trg}(t)$ (80% with 32000 data points) and $\hat{X}_{tst}(t)$ (20% with 8000 data points) sets that are treated as the training and testing sets for a simulation based validation. Then we artificially inject sensor faults into the testing set as described above. Simulation cases include: (1) adding drift from 1 mile per

Table 3.1 Anomaly detection results with one-sensor-fault simulations

Noise Type	Data Type	Method	Severity									
			1	2	3	4	5	6	7	8	9	10
Drift(mile/hour)	S	STPN	0/1	0/1	0/1	0/1	1/1	1/1	1/1	1/1	1/1	1/1
		RCA	0/1	0/1	0/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1
	S+V	STPN	0/1	0/1	0/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1
		RCA	0/1	0/1	0/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1
	S+V+O	AEVL	0/1	0/1	0/1	0/1	0/1	0/1	0/1	0/1	0/1	1/1
Gaussian Noise(mile/hour)	S	STPN	0/1	0/1	0/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1
		RCA	0/1	0/1	0/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1
	S+V	STPN	0/1	0/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1
		RCA	0/1	0/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1	1/1
	S+V+O	AEVL	0/1	0/1	0/1	0/1	0/1	0/1	0/1	0/1	0/1	1/1

hour to 10 miles per hour to one sensor (sensor 3, $\mathcal{S}3$), two sensors ($\mathcal{S}3$ & $\mathcal{S}7$), and five sensors ($\mathcal{S}1$, $\mathcal{S}2$, $\mathcal{S}3$, $\mathcal{S}6$, and $\mathcal{S}7$) respectively; (2) adding Gaussian noise with standard deviation from 1 to 10 to one sensor ($\mathcal{S}3$), two sensors ($\mathcal{S}3$ & $\mathcal{S}7$), and five sensors ($\mathcal{S}1$, $\mathcal{S}2$, $\mathcal{S}3$, $\mathcal{S}6$, and $\mathcal{S}7$) respectively. The cases are tested with AEVL, off-line (STPN) and online (RCA) approaches as presented in Section 3. Note that we call the STPN+RBM based online method as RCA as it essentially uses a root-cause analysis (RCA) approach to solve the sensor fault detection and isolation problem.

The results are summarized in Table 3.1–3.3 where the drift and noise variance levels are treated as severity levels. The simulation results show that 2 dimensional models have higher sensitivity than 1 dimensional models when using STPN and RCA and both can isolate the faulty sensors with a higher sensitivity. The benchmark method AEVL does not perform as well despite using 3 sensor modalities.

3.4.2 Sensor Fault Detection with Real Data

The real use case involves the original training data and testing data ($X_{trg}(t)$ and $X_{tst}(t)$) that are collected for the same sensor network every 20s by Wavetronix LLC. in 2016. Using detail manual investigation, we find that sensor 6 is anomalous during the testing period and hence used as the ground truth. In the benchmark method, we can compute the AEVL for each sensor and

Table 3.2 Anomaly detection results with two-sensors-fault simulations

Noise Type	Data Type	Method	Severity									
			1	2	3	4	5	6	7	8	9	10
Drift(mile/hour)	S	STPN	0/2	0/2	0/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2
		RCA	0/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2
	S+V	STPN	0/2	0/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2
		RCA	1/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2
	S+V+O	AEVL	0/2	0/2	0/2	0/2	0/2	0/2	0/2	0/2	0/2	2/2
Gaussian Noise(mile/hour)	S	STPN	0/2	0/2	0/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2
		RCA	0/2	1/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2
	S+V	STPN	0/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2
		RCA	1/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2
	S+V+O	AEVL	0/2	0/2	0/2	0/2	0/2	0/2	0/2	0/2	0/2	2/2

Table 3.3 Anomaly detection results with five-sensors-fault simulations

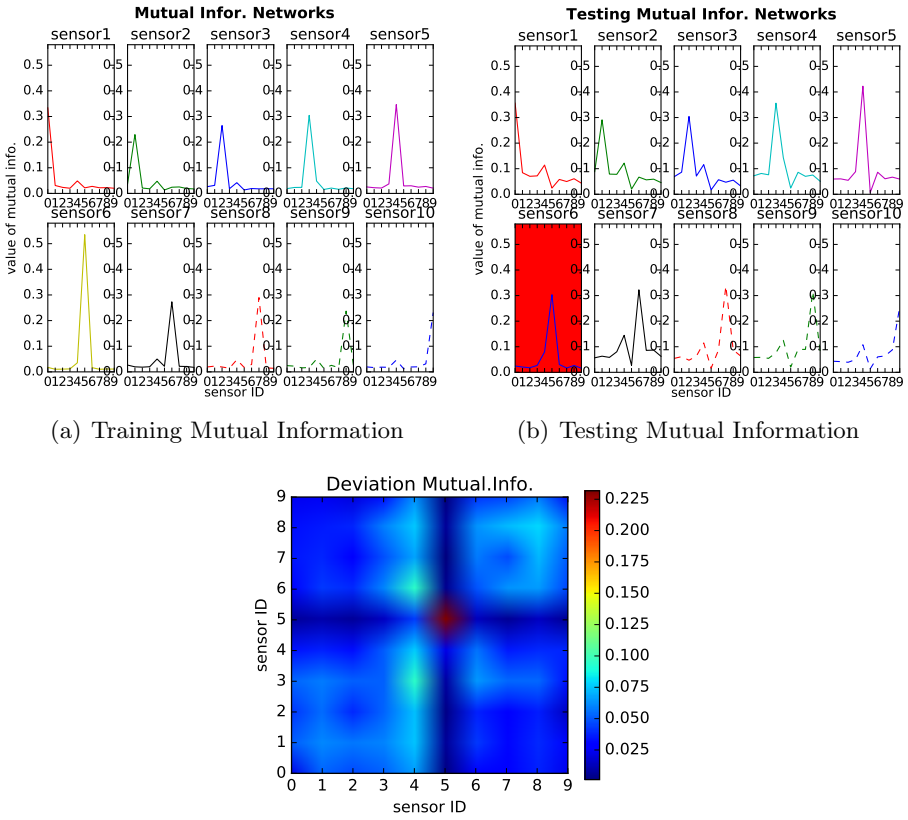
Noise Type	Data Type	Method	Severity									
			1	2	3	4	5	6	7	8	9	10
Drift(mile/hour)	S	STPN	0/5	2/5	2/5	2/5	3/5	3/5	3/5	3/5	5/5	5/5
		RCA	0/5	2/5	2/5	2/5	3/5	3/5	3/5	3/5	4/5	4/5
	S+V	STPN	0/5	2/5	2/5	3/5	3/5	3/5	5/5	5/5	5/5	5/5
		RCA	2/5	2/5	3/5	3/5	3/5	3/5	4/5	4/5	4/5	5/5
	S+V+O	AEVL	0/5	0/5	0/5	0/5	0/5	0/5	0/5	0/5	0/5	5/5
Gaussian Noise(mile/hour)	S	STPN	0/5	0/5	0/5	3/5	3/5	4/5	4/5	4/5	5/5	5/5
		RCA	2/5	2/5	2/5	2/5	3/5	3/5	3/5	4/5	4/5	4/5
	S+V	STPN	1/5	4/5	4/5	4/5	4/5	5/5	5/5	5/5	5/5	5/5
		RCA	2/5	3/5	3/5	3/5	3/5	4/5	4/5	4/5	4/5	5/5
	S+V+O	AEVL	0/5	0/5	0/5	0/5	0/5	0/5	0/5	0/5	0/5	5/5

S:speed, V: volume, O: occupancy. The results in above tables (table 3.1–3.3) are represented in m/n, where m represents the detected anomaly sensor(s) and n denotes the sensor(s) are labeled anomaly in the sensor network. The severity levels correspond to the different levels of bias and noise variance synthetically added to the test data.

Table 3.4 Results of RCA and AEVL with real data

Methodology	Data Type	Probability of detected faults (ranked)
RCA	Speed	**S6=0.512** S1=0.237 S2=0.151
	Speed+Volume	**S6=0.725** S10=0.210
AEVL	Speed+Volume+Occupancy	S10=0.061 S8=0.051 S7=0.031 S5=0.029

S1 to S10 represents the sensor ID



(a) Training Mutual Information

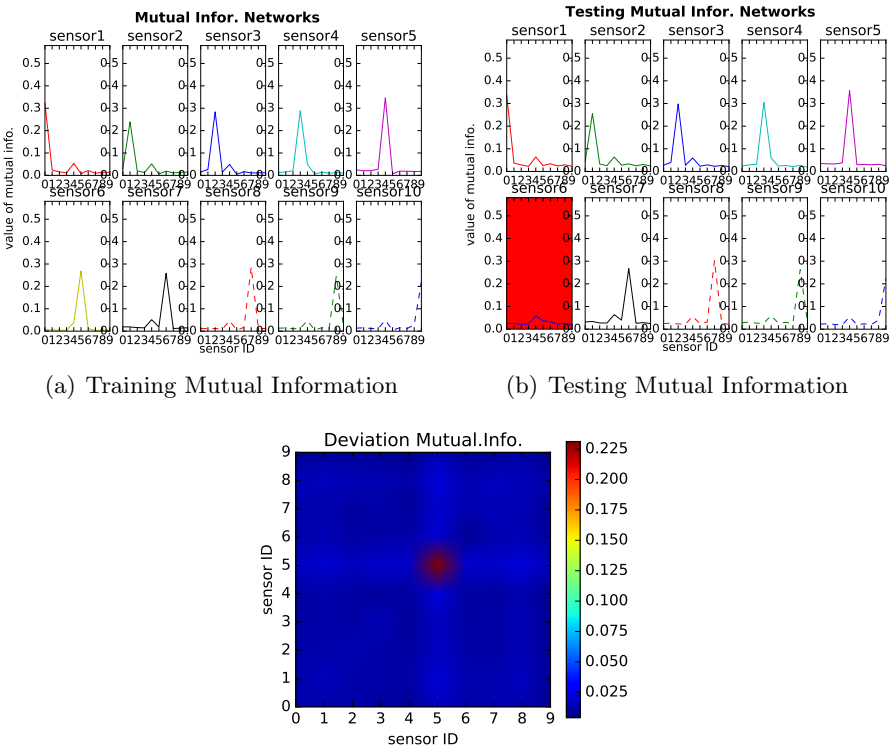
(b) Testing Mutual Information

(c) Detected Result

Figure 3.6 *Labeled sensor health monitoring using STPN with vehicle speed data, where sensor 6 has been detected as an anomaly sensor (matched with the labeled data) and the result is directly reported in image c with some ambiguity.*

use this value to compare with the length range (10 feet to 75 feet as shown in section 3.3.2) and compute the number of error occurrences (denoted by e). Then, the AEVL differential error will be: $\delta E_r^i = \frac{e_{trg}^i}{\#X_{trg}^i} - \frac{e_{tst}^i}{\#X_{tst}^i}$ for the i^{th} sensor, and $\#X_{trg}^i$ represents the total number of data points in the training data. Then, δE_r^i is the error rate or anomaly score in this case for each sensor in the network.

For the off-line method (i.e., STPN), the symbol sequences, S^1, S^2, ..., S^{10}, are generated by symbolizing $X(t)$ (both $X_{trg}(t)$ and $X_{tst}(t)$) via SSD and MBD partitioning techniques. And then the mutual information matrices $\underline{\Lambda}_{trg}$ and $\underline{\Lambda}_{tst}$ can be obtained. To identify the anomalous sensor, deviation in mutual information matrices, $\Delta\underline{\Lambda} = \underline{\Lambda}_{trg} - \underline{\Lambda}_{tst}$ can be computed as visualized in



(a) Training Mutual Information

(b) Testing Mutual Information

(c) Detected Result

Figure 3.7 *Labeled sensor health monitoring using STPN with vehicle speed and volume data, where the labeled fault sensor 6 has been detected with less ambiguity as in Fig.3.6.*

Figs. 3.6-3.7. As shown in Fig. 3.6, we can conclude that the faulty sensor 6 can be identified by STPN only using the speed data. However, there may be ambiguities and false alarms as evidenced visually by the deviation matrix. On the other hand, when using both speed and volume, the off-line STPN method can detect the fault more accurately without any significant ambiguity as seen in Fig. 3.7.

The on-line method adopts short sequences U, where $U \subset (S^1, S^2, ..., S^{10})$ to form the inference metric and then implements root-cause analysis to identify the anomalous sensor(s). Table 3.4 presents the ranking of possibly faulty sensors along with their respective anomaly scores. The on-line approach can correctly isolate the faulty sensor (which shows the highest anomaly score), while the AEVL method can not identify the faulty sensor despite using 3 different sensing modalities.

Remark. *Although the proposed online detection technique is susceptible to false alarms, it plays a critical role in the entire traffic sensor health monitoring framework. This can be used for getting early indication of possible sensor anomalies and failures in traffic systems. They can be manually verified only if a potentially anomalous sensor is crucial such as being safety-critical. In this case, a certain level of false alarm can be acceptable. The online detection technique is complemented by the off-line method which can verify the alarms from the online technique using batch processing.*

3.5 Summary

By applying the concept of spatiotemporal pattern network, this work proposes two ways (on-line and off-line) to monitor sensor health via graphical modeling of sensor network data. Both approaches are designed to process large-scale time-series data in sensor networks with advantages in: (1) extracting spatiotemporal features to discover relationships among sensors, (2) detecting anomaly in an off-line manner by computing an information theoretic metric, (3) monitoring and localizing anomalous sensor in real-time by computing an inference based metric.

Based on the results, it can be concluded that: (i) compared with the benchmark AEVL method, off-line and on-line methods can isolate the anomalous sensor more accurately and can be very effective for different types of sensor anomalies such as bias or drift and increased noise variance, (ii) 2D model using both speed and volume data can distinguish the anomalous sensor more clearly than the 1D model. The possible reason is that, regional congestion may affect the 1D result since the speed is very low at that point, while the relationship among the sensors remain preserved while using both the speed and volume information. Also, while the proposed off-line method is more stable and sensitive, the on-line approach is fast, i.e., low time-to-detect (suitable for real-time application) but may have more false alarms.

CHAPTER 4. A DATA-DRIVEN APPROACH TOWARDS INTEGRATION OF MICROCLIMATE CONDITIONS FOR PREDICTING BUILDING ENERGY PERFORMANCE

4.1 Introduction

Energy consumption in buildings is a major part of the overall energy usage in the United States and across the world. Ambient weather conditions are the main factors affecting the heat exchange from surroundings to the inside of building. Generally, the building surrounding environment information is considered to be from meteorological weather data Meek and Hatfield (1994) or a local micro weather station and is used in various building energy simulation (BES) tools for energy consumption or energy costs analysis Swan and Ugursal (2009). Given the fact that the microclimatic conditions around the building are affected by the neighboring buildings, vegetation, air flow patterns etc. These factors also affect the energy performance of the building. The local weather station climatic data, since recorded in the far-field, may not account the local climatic information around the building. Hence, it is important to include microclimate information to have a better understanding of the building energy performance. The significance and the use of microclimatic information are shown for various studies, such as in urban heat islands for peak building energy use and energy savings Bhiwapurkar (2015). Further, microclimatic conditions in urban environment can also significantly affect urban energy management, important in development of cities or certification of buildings based on their energy performance, such as LEED Scofield and Oberlin (2009), Energy Star Boyd et al. (2008) etc. The important microclimatic variables are average surrounding temperature, wind, humidity, solar radiation, and vegetation. These variables are usually monitored around the building. For such monitoring, a spatiotemporal relationship with the local weather data can be constructed as a model to predict the microclimatic condition.

Various efforts are made in the area of integrating the microclimatic conditions for predicting the energy performance of buildings, a few are listed below: 1) A procedure for coupling microclimate model ENVI-met with EnergyPlus for quantitative analysis of building energy performance by Yang et al. (2012). 2) A simulation platform relying on both computational fluid dynamics (CFD) and themoradiative simulation codes by Bouyer et al. (2011) to measure the impact of urbanization and planning strategies on an energy demand. In the work they also did sensitivity analysis for finding important microclimatic variables affecting building energy consumption. 3) Coupling CFD and BES are shown by Dorer et al. (2013) for analyzing the impact of urban microclimate (UMC) on space heating and cooling energy demand. In the work by Yuming et.al Sun et al. (2014), a framework is proposed to compare the microclimate variables computed from standard model in BES programs to those with mesoscale model which takes into account more detailed specifications of urban form. Further, the vegetation effect of the energy performance of the building is also studied in detail by Huang et al. (1987). The occupancy which act as a important parameter in building design, in association with microclimate can be looked in the work by Kalvelage et al. (2015).

In this study, we consider a real-time monitored and operated building at Ottumwa, Iowa. The availability of data pertaining to the key microclimatic variables for the building makes it possible to apply data driven approaches to have a deeper understanding of patterns and process in building energy systems. This motivates us to use data driven approaches for microclimatic modeling and analyze its impact on the energy performance. The main goal of this work is to present a proof of concept to analyze the impact of microclimatic information on the energy consumption of buildings using data-driven approaches Jiang et al. (2017); Sarkar and Srivastav (2016).

Contributions: The major contributions of this work are: (i) the efficacy of using predicted microclimate information for energy consumption prediction as compared to the ambient weather data. (ii) the use of data-driven models in predicting microclimate variables and using those predictions for energy consumption prediction based on regression models. (iii) validating the prediction and regression models using the real data of Interlock house located in Ottumwa, Iowa.

Figure-4.1 shows the pictorial representation of the testbed considered for the data collection in the study.

Figure 4.1 *Interlock house located at the Ottumwa, Iowa, used as a testbed in the study. It is equipped with energy and microclimate monitoring sensors.*

4.2 Data-Driven Methodologies

Building energy systems consist of different subsystems interacting at different temporal and spatial scale, and are therefore complex to simulate. It is important to account for these temporal and spatial dependencies of system properties when modeling building systems. With the technological advances and data driven learning methods, it is easy to use the available data, infer patterns, and information underlying the process. In this study, two different data driven approaches have

been used: 1) Spatiotemporal pattern networks (STPNs) 2) Neural Networks (NNs) to predict the microclimate conditions and then use it for predicting energy performance in buildings. A brief summary of details of both the methods are listed in below sections.

4.2.1 STPN Framework for Multi-variables

In this work we propose a model with multiple time-series inputs and a single prediction target. Suppose $X = \{X^{\mathbb{A}}(t), t \in \mathbb{N}, \mathbb{A} = 1, 2, .., n\}$ represents the multivariate time-series data from city weather station, n is the number of time-series and $Y = Y(t), t \in \mathbb{N}$ denotes the microclimatic information (here is temperature) collected by microclimate station. And for the time-series X, Y we have symbol sequences $\sigma = \{\sigma^{\mathbb{A}}(t)\}$ and $\mu = \{\mu(t)\}$ after data processing and partitioning. To generate the dependencies we form a joint symbol sequence for X, $\sigma^{\mathbb{J}} = \sigma^1(t) \oplus ... \oplus \sigma^n(t)$. Here, the joint symbol space is created by summing the individual symbol spaces directly. For example, $\sigma^a \oplus \sigma^b$ is the product space of σ^a and σ^b. Then based on the previous definition of STPN we can get the state sequences Φ and Ψ respect to $\sigma^{\mathbb{J}}$ and μ. The training step of STPN is to compute the transition matrix $\omega(\Phi, \mu)$ from the states in Φ to the symbols in μ using a frequentist's technique(e.g. counting the number of occurrences). For example, we can compute the probability of the state Φ_m to the symbol μ_i by $Pr(\phi_m, \mu_i) = N_{mi}/N_m$, where N_{mi} is the number of times that the symbol $\mu_i \in \mu$ is emanated after the state $\phi \in \Phi$, and here $N_m \triangleq \sum_{i=1}^{n} N_{mi}$. In the model, the true values represented by μ can be computed from the training data, and noted as $E(X|\mu = n), n = 1, 2, ...$ In the prediction model, we only have the time-series X and the microclimatic temperature time-series $\hat{Y}$ is assumed unknown and used for the verification. And similarly, we can get the symbol sequences $\hat{\sigma}$ and joint symbol sequence σ^J as well as state sequences. Thus, the microclimatic temperature Y_t is obtained as:

$$Y_t = \sum_{i=1}^{n} Pr(\mu|\hat{\phi}(t)) \times E(Y|\mu = i) \tag{4.1}$$

A schematic of the procedure of STPN is shown in Fig. 4.2

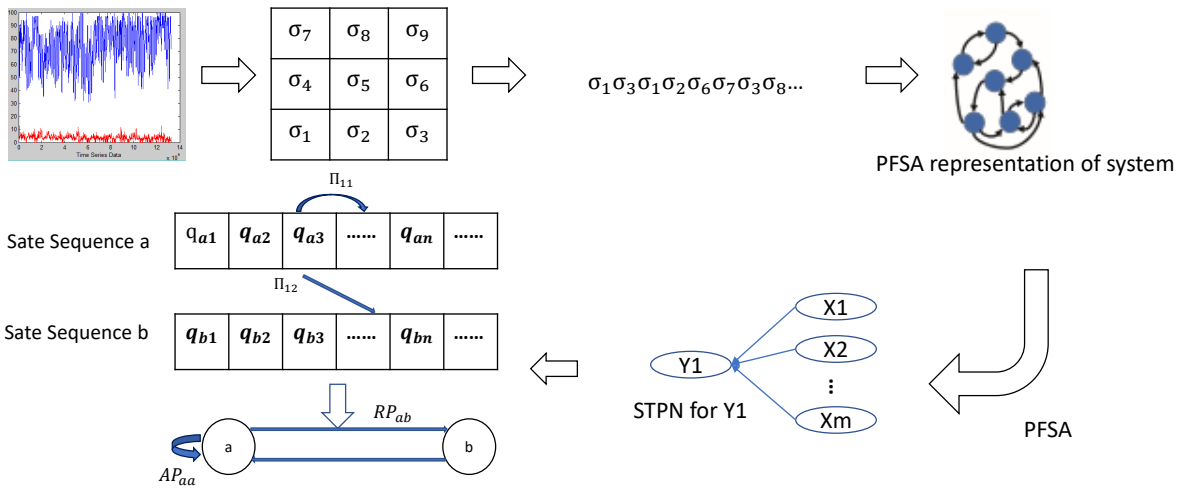


Figure 4.2 *Formulation of STPN with multiple time-series (nodes in graphical modeling). It extracts atomic patterns (AP) and relational patterns (RP) with D-Markov machine and xD-Markov machine respectively and the depth D=1.*

4.2.2 Neural Networks

Neural networks (NNs) are inspired by the biological operation of the specialized cells called neurons. In a human body a neuron is a cell which receive several inputs activated from the outside process. Based on the activation level, the neuron gets stimulated and produces the outputs. Further, each input and output path of the neuron can have their own strength and can differ from one another. The neural network tries to replicate this by a set of weighted graphs and a neuron is represented as a node. The node receive inputs, and processes their sum with its *activations function* ϕ and passes the results to the nodes further in the graph. Mathematically this can be written as

$$\phi(\mathbf{w^T a}) = \frac{\mathbf{1}}{\mathbf{1} + \mathbf{exp}(-\mathbf{w^T a})} \tag{4.2}$$

The sigmoid activation function is used at each node in the network. A schematic of the neural network is shown in Fig-4.3. The nodes are chained together to form the hidden layer in the network, there can be multiple set of hidden layer which passes the output as a input to the next hidden layer to reach the final output. The goal is to train a neural network based on the data so

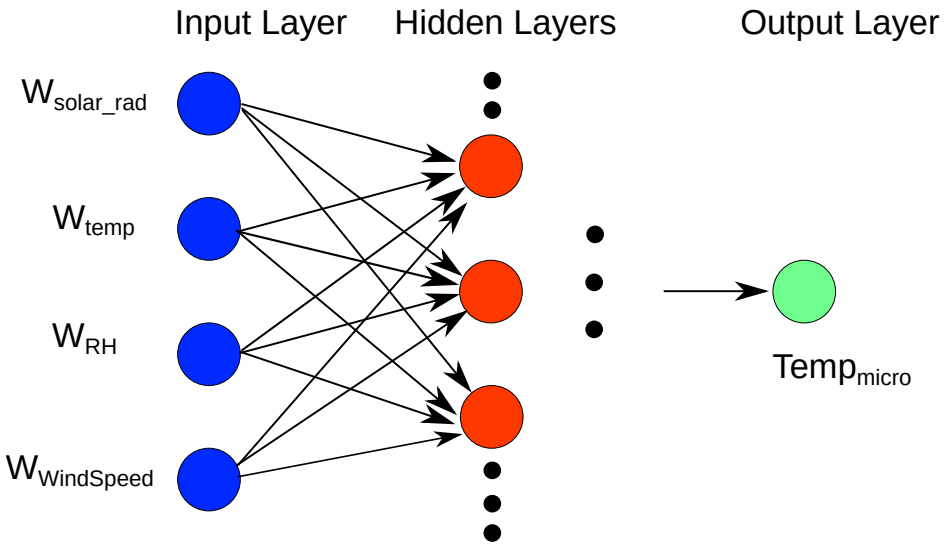


Figure 4.3 *Schematic of a Neural network trained in the current study for predicting the microclimate variables based on the ambient weather data*

that it can be used to predict for a set of unseen input data.

Recent success of deep learning has demonstrated the enormous possibilities of wider and deeper neural networks for learning complex functions without feature hand-crafting Lore et al. (2015). However, in the present work it is difficult to build a large neural network due to the small volume of data. The regular neural network developed here was trained using time-series data of the microclimate and weather station data sampled at the same time-interval. The network can be used for classification as well as regression and for the current study we develop regression model, for which we define the loss function as follows,

$$L(\mathbf{w}) = \sum_{\mathbf{i}} \mathrm{f}(\mathbf{x_i}, \mathbf{w}) - \mathbf{y_i})^{\mathbf{2}} \tag{4.3}$$

where, $\mathrm{f}(\mathbf{x_i}, \mathbf{w})$ is the regression function and we must minimize $L(\mathbf{w})$. This equation has the closed form solution for general least squares, but here we use the gradient based method, ADAM Kingma and Ba (2014) to minimize the loss. In general, the gradient of the loss gradient $\nabla_{\mathbf{w}} L(\mathbf{w})$ is to minimize the overall error on the training data. The first step is to derive the loss with respect to a particular weight $w_{j \to k}$ (This is the weight of the edge connecting node j to node k) for the

general setting we write this as follows:

$$\begin{aligned}\frac{\partial}{\partial w_{j\to k}}L(\mathbf{w}) &= \frac{\partial}{\partial w_{j\to k}}\sum_i(\mathrm{f}(\mathbf{x_i},\mathbf{w})-\mathbf{y_i})^{\mathbf{2}}\\ &= \sum_i\frac{\partial}{\partial w_{j\to k}}(\mathrm{f}(\mathbf{x_i},\mathbf{w})-\mathbf{y_i})^{\mathbf{2}}\\ &= \sum_i 2(\mathrm{f}(\mathbf{x_i},\mathbf{w})-y_i)\frac{\partial}{\partial w_{j\to k}}\mathrm{f}(\mathbf{x_i},\mathbf{w})\end{aligned} \tag{4.4}$$

This general process is modified in ADAM by the stochastic gradient finding method (more details are provided in Kingma and Ba (2014)). Following the discussed procedure, the network is trained to predict the microclimate data as a dependent variable by using the weather station data as input. The trained model is then used to input the conditions to predict the energy performance of the building. We use the MSE (mean square error) to calculate the accuracy of the model, as a widely accepted metric for regression problems. The quality and quantity affect the performance and with more time-points the MSE can be reduced further for the current problem. We then use the model for predicting the microclimatic data, which then can be used for energy predictions.

4.3 Test Schematic & Data Preparation and Model Setup

4.3.1 Test Schematic

The testing building, interlock house, located in Ottumwa, Iowa as shown in Fig. 4.1 is equipped with sensors for thermal performance and energy consumption monitoring and collecting the surrounding microclimatic information (i.e., temperature, humidity, wind speed etc.). A schematic of the house is shown in Fig 4.4. It is a solar energy based building close to the Honey Creek Resort State Park, which has the best geographic location with the longest time for sunshine. The heating load is calculated through the heating gain model:

$$H_{load} = (T_{in} - T_{out}) \cdot C_w \cdot V_w \cdot \rho_w \tag{4.5}$$

where H_{load} is the heating load (energy consumption in winter), T_{in} is the temperature of the flowing in water, T_{out} is the temperature of the flowing out water, C_w is the specific heat of

water, V_w is the volume and ρ_w is the density of water. Given the microclimate data of interlock house and the associated ambient weather data, a data driven approach is proposed to predict the microclimatic conditions (temperature). The predicted temperature is in turn used for energy consumption prediction based on a regression analysis.

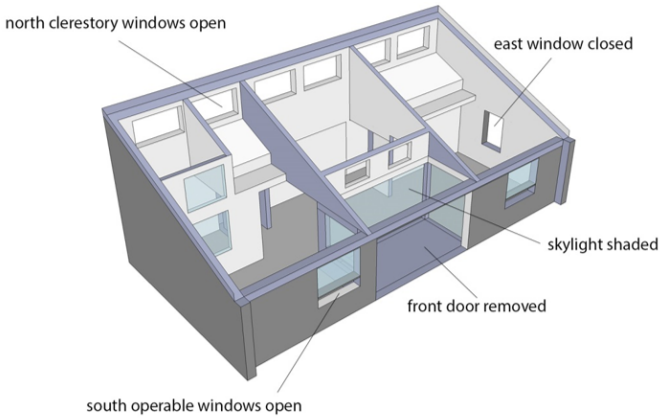


Figure 4.4 *The building used in the study (Interlock house), the yearly energy data and microclimatic data is recorded for the building.*

4.3.2 Data Preparation and Model Setup

The workflow chart for the proposed data-driven energy analysis is shown in Fig. 4.5. Typical meteorological data (TMY3) from Ottumwa city is considered for the weather station data and the key variables used for the regression analysis are outside air temperature (T_a), relative humidity (RH), global horizontal irradiation (E_e), and wind speed (V_w). For the simulation study, the weather station data X from the months of July, August and September of year 2015 is considered as the summer verification season (X_1) and from December 2015 to February 2016 are the winter verification season (X_2). And the same timestamp for interlock house microclimatic variables Y. Those data sets are 1-minute interval data and there are 129,500 data points for each season being used for the analysis. The data is further split into training ($X_{tr} = 85\%X, Y_{tr} = 85\%Y$) and validation ($X_{ts} = 15\%X, Y_{ts} = 15\%Y$) for use by both the data-driven approaches.

The simplified model for both data-driven methodologies are:

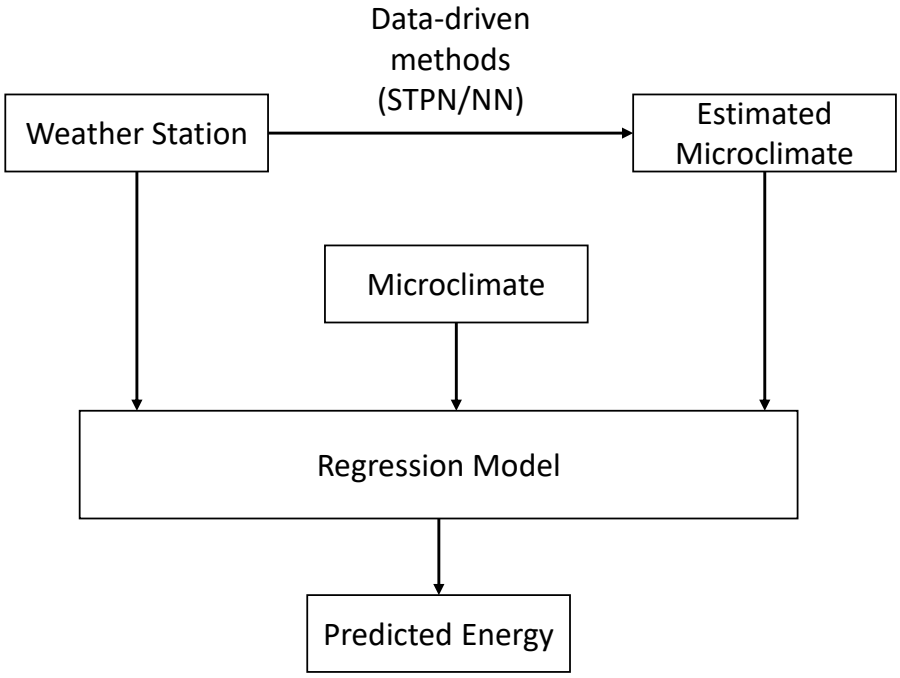


Figure 4.5 *Workflow procedure for incorporating the microclimatic conditions using the data-driven approach.*

$$Y_{temp} = f(X(t)|X(\tau : t-1), Y_{temp}(\tau : t-1)) \tag{4.6}$$

or:

$$\begin{aligned} T_{micro}(t) = f(&T_a(t), RH(t), E_e(t), V_w(t)|T_a(\tau : t-1), \\ &RH(\tau : t-1), E_e(\tau : t-1), V_w(\tau : t-1), \\ &T_{micro}(\tau : t-1)) \end{aligned} \tag{4.7}$$

Where t is the target time and τ is the training start time.

The predictions of microclimate variables obtained using the data driven approaches (STPN/NN) are then used for energy prediction analysis. For energy prediction, we use autoregressive-moving-average (ARMAX) Yang et al. (1996) regression model using the four regression variables (RH,T_a,E_e,V_w).

These models build the relationship between the inputs and the outputs of the system using measurements and also helps in understanding the energy characteristics of the building.

4.4 Results and Discussion

In this work, the proposed STPN and NN models are used for microclimatic temperature prediction. Further, these prediction results are used for energy prediction analysis using regression method. We conduct validation in two different seasons, summer and winter to analyze the affect of climatic conditions on the prediction performance. In the following, this study provides the detailed results of using the proposed methods as well as comparisons.

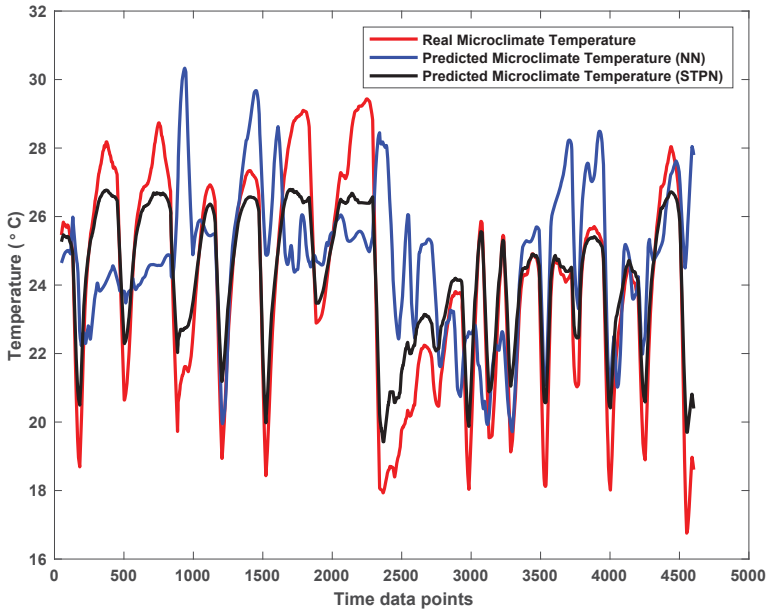


Figure 4.6 *Comparisons of real microclimate temperature and the predicted microclimate temperature using STPN and Neural Networks methods for summer season.*

4.4.1 Summer Season

The procedure for using STPN and NN methods are detailed in Section 4.2. For the summer season (mainly focus on cooling), the data from months of July, August, September of 2015 are considered. Key variables from the ambient weather data are used for prediction of microclimate

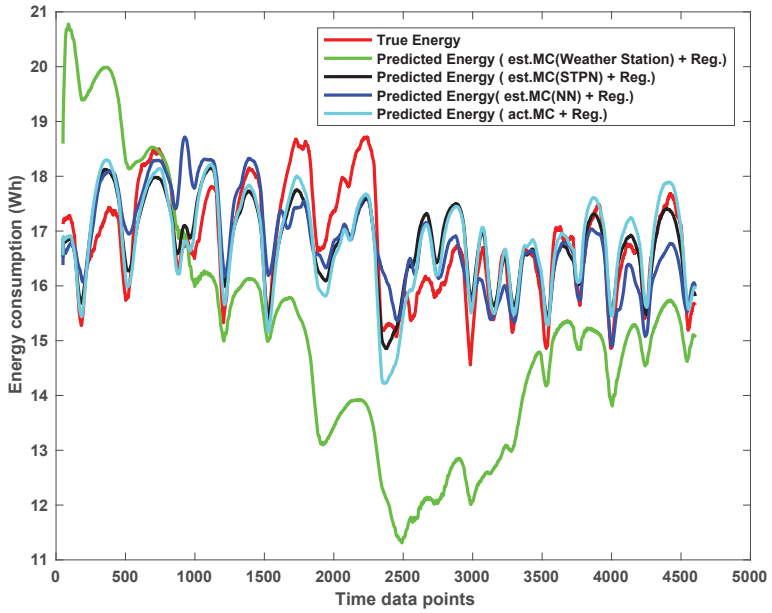


Figure 4.7 *Comparison of energy consumption prediction using predictions from data–driven approaches, ambient weather data, actual microclimate data and true energy consumption. Here, est.MC represents predicted microclimate data, Reg. means the regression model and act.MC means the actual microclimate data.*

temperature. The testbed has outdoor AC units which are used for providing cooling during the summer season. The data has been preprocessed to filter any outliers (i.e. manually shut off) in the data set which affects the prediction results. The temperature prediction results using both the methods in comparison to actual temperature is shown in Fig. 4.6. The prediction results using both the methods captures the trend associated with the data and STPN performs better compared to NN. For the case of STPN, the error in the peaks can be improved by partitioning the data into finer bins. Typically, NNs perform well while being trained with a large data set (by avoiding overfitting). Therefore, in this case the slightly lesser accuracy compared to STPN can be attributed to the data quantity as well as lack of hyperparameter tuning.

These temperature prediction results are used in the regression model for energy prediction. A comparison plot of energy prediction using the prediction results from the proposed methods along with energy predicted from weather station data is shown in Fig. 4.7. The results indicate that the energy prediction from the data-driven methods has higher accuracy (i.e. fit better) compared

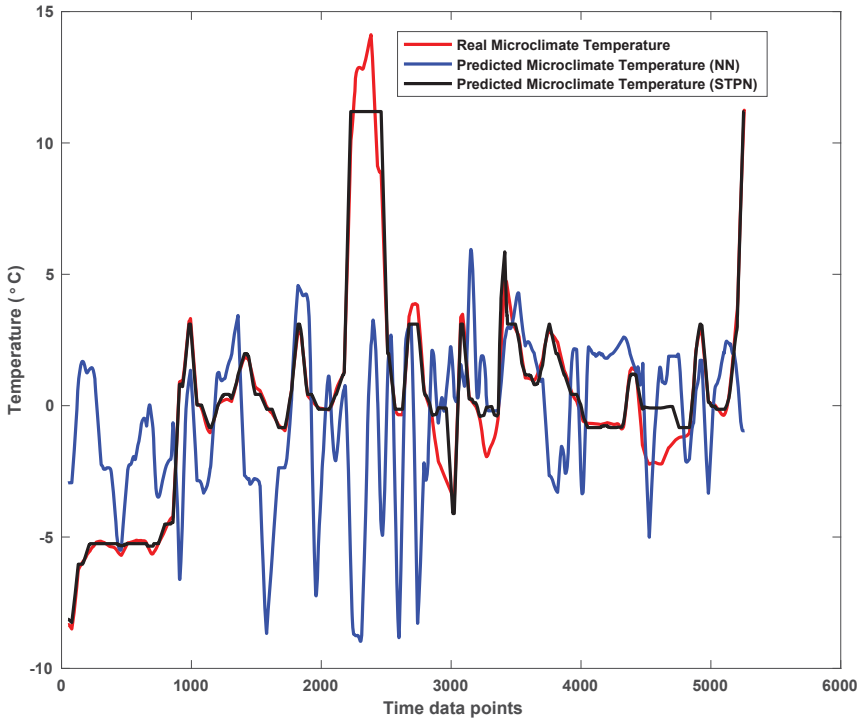


Figure 4.8 *Comparisons of real microclimate temperature and the predicted microclimate temperature using STPN and Neural Networks methods for winter season.*

to the energy predicted from ambient weather station data. This clearly signifies the importance of microclimate information on the building energy consumption prediction analysis. This study will be further extended in future to include building specific variables (e.g., building architecture, materials, vegetation around the building) for improving the prediction performance.

4.4.2 Winter Season

For the winter season the months of December 2015, Januray & February of 2016 are considered. The microclimate temperature prediction results are shown in Fig. 4.8. The prediction result of STPN method has improved performance over the NN method. Similarly the energy consumption comparison result is shown in Fig. 4.9. The energy consumption results also suggest that microclimate information is crucial in predicting the building heating energy consumption. The heating energy consumption is mainly affected by the heat transfer between inside and outside of

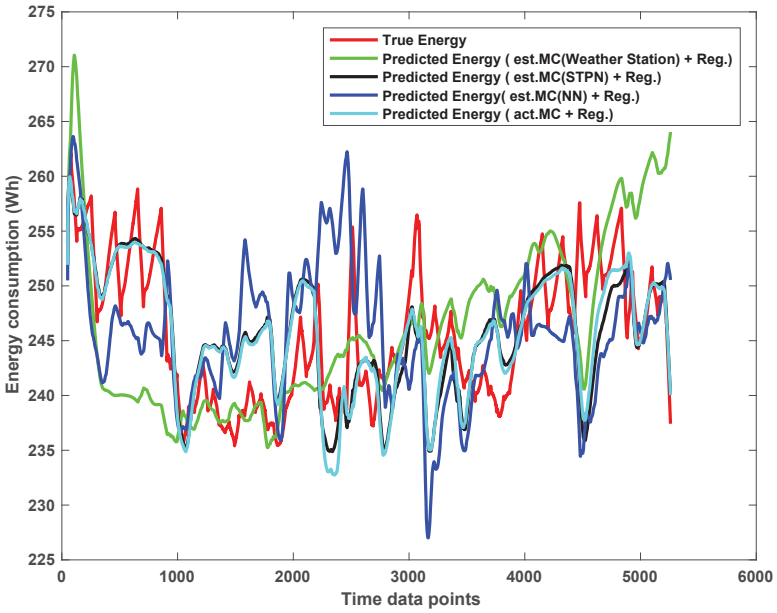


Figure 4.9 *Comparison of energy consumption prediction using predictions from data–driven approaches, ambient weather data, actual microclimate data and true energy consumption.*

the building. From the Fig. 4.8 we can see the temperature in winter changes in a larger range, which is more complicated for data-driven model to learn the patterns. The results can be further improved by including additional variables which captures the heat loss from the building which in turn affects the heating energy consumption.

For a multivariate time-series, the proposed STPN framework captures the informative patterns, which results in predicting microclimate data. The case studies in this study show that, although microclimatic information is not always available for every building, data-driven techniques help in obtaining a better energy consumption prediction. Also, the data quality and quantity can significantly improve the microclimatic prediction and hence the energy prediction results. Overall, both data-driven methodologies contribute a great effects on the prediction of building energy consumption, especially, STPN method performs better in comparison with NN method due to the limited data points. Using variants of NN for prediction analysis needs to be investigated further. Energy prediction comparison results are shown in Table 4.1. The mean square error using

predictions based on STPN method is close to the real microclimate data, which suggests that the data-driven models are sufficient for predicting microclimate information which replaces the need for sensing cost in obtaining such information.

Table 4.1 Comparison of energy prediction using different weather conditions

Mean square error with different types of weather data				
Season	Ambient weather station data	Real microclimatic data	Predicted microclimatic data using data-driven models	
Summer	7.33	**0.54**	STPN	**0.56**
			NNs	0.70
Winter	88.07	**41.45**	STPN	**44.04**
			NNs	71.64

4.5 Conclusion

This work presents a proof of concept study to demonstrate the efficacy of using microclimate information for building energy consumption prediction as compared to ambient weather data. In this regard, data-driven approaches can be extremely effective for prediction of microclimate variables and hence, can help avoid expensive microclimate sensor deployment. Quantitative results (using energy prediction accuracy metrics such as mean square error) show that the energy prediction from the data-driven approaches are sufficiently close to the actual microclimate data. These results illustrate that the predicted microclimate information based on data-driven approaches is a reliable and low-cost way to predict building energy consumption.

CHAPTER 5. SUMMARY AND CONCLUSION

This thesis has demonstrated that the proposed graphical modeling methodology with domain knowledge holds enormous potential in improving the management performance of cyber-physical systems, health monitoring, and estimation. Specifically, spatiotemporal pattern networks (STPNs) are capable of modeling complicated nonlinearities time-series data as discussed in chapter 2. In chapter 3, this work constructs the off-line detection model based on STPN and combines STPN with Restricted Boltzmann machines for the online model for identifying the source of anomalies. The effectiveness of applying graphical modeling for traffic sensor network health monitoring has been validated by comparing with the previous works as shown in the results section in chapter 3. Additionally, applying STPN for multivariate model to exploit the inherent correlations between local weather station and microclimate conditions is significantly helpful for the applications where a scalable and accurate estimation of building energy consumption needs to be made based on local weather data. A representative application integrating microclimate conditions for predicting building energy consumption has been studied in this thesis in chapter 4.

In contrast to other machine learning methods, e.g. neural networks, STPN for multivariate model has a signicant potential in providing scalable and effective solutions to address engineering problems involving spatiotemporal data. The original motivation behind proposing graphical modeling methodology for the health monitoring and performance prediction in cyber-physical systems is in fact to pursue the cost-effective and robust techniques as mentioned in chapter 1. The future work will focus on (1) visualizing the result through an APP for customer view and (2) using deep learning based modeling techniques such as long short-term memory for other engineering applications, such as cyber-agriculture systems.

BIBLIOGRAPHY

(2016). Federal size regulations for commercial motor vehicles. *U.S. Department of Transportation Federal Highway Administration*.

(2016). Wavetronix smartsensor hd user guide. *Wavetronix*.

Amasyali, K. and El-Gohary, N. M. (2018). A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews*, 81:1192–1205.

Baheti, R. and Gill, H. (2011). Cyber-physical systems. *The impact of control technology*, 12:161–166.

Basir, O. and Yuan, X. (2007). Engine fault diagnosis based on multi-sensor information fusion using dempster–shafer evidence theory. *Information Fusion*, 8(4):379–386.

Bhiwapurkar, P. (2015). Urban climate change impacts on building heating and cooling energy demand. *of Architectural Research*, page 270.

Bhuiyan, M. Z. A., Wang, G., and Wu, J. (2009). Target tracking with monitor and backup sensors in wireless sensor networks. In *Computer Communications and Networks, 2009. ICCCN 2009. Proceedings of 18th Internatonal Conference on*, pages 1–6. IEEE.

Bouyer, J., Inard, C., and Musy, M. (2011). Microclimatic coupling as a solution to improve building energy simulation in an urban context. *Energy and Buildings*, 43(7):1549–1559.

Boyd, G., Dutrow, E., and Tunnessen, W. (2008). The evolution of the energy star® energy performance indicator for benchmarking industrial plant manufacturing energy use. *Journal of cleaner production*, 16(6):709–715.

Brimacombe, J. M., Wilson, D. R., Hodgson, A. J., Ho, K. C., and Anglin, C. (2009). Effect of calibration method on tekscan sensor accuracy. *Journal of biomechanical engineering*, 131(3):034503.

Butte, A. J. and Kohane, I. S. (1999). Mutual information relevance networks: functional genomic clustering using pairwise entropy measurements. In *Biocomputing 2000*, pages 418–429. World Scientific.

Chen, C., Petty, K., Skabardonis, A., Varaiya, P., and Jia, Z. (2001). Freeway performance measurement system: mining loop detector data. *Transportation Research Record: Journal of the Transportation Research Board*, (1748):96–102.

Dailey, D. J. (1999). A statistical algorithm for estimating speed from single loop volume and occupancy measurements. *Transportation Research Part B: Methodological*, 33(5):313–322.

Dorer, V., Allegrini, J., Orehounig, K., Moonen, P., Upadhyay, G., Kämpf, J., and Carmeliet, J. (2013). Modelling the urban microclimate and its impact on the energy demand of buildings and building clusters. *Proceedings of BS*, 2013:3483–3489.

Douglas, D. H. and Peucker, T. K. (2011). Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Classics in Cartography: Reflections on Influential Articles from Cartographica*, pages 15–28.

Fiore, U., Palmieri, F., Castiglione, A., and De Santis, A. (2013). Network anomaly detection with the restricted boltzmann machine. *Neurocomputing*, 122:13–23.

González, P. A. and Zamarreno, J. M. (2005). Prediction of hourly energy consumption in buildings based on a feedback artificial neural network. *Energy and Buildings*, 37(6):595–601.

Guan, X., Xu, Z., and Jia, Q.-S. (2010). Energy-efficient buildings facilitated by microgrid. *IEEE Transactions on smart grid*, 1(3):243–252.

Harris, T., Gamlyn, L., Smith, P., MacIntyre, J., Brason, A., Palmer, R., Smith, H., and Slater, A. (1995). 'neural-maine': intelligent on-line multiple sensor diagnostics for steam turbines in power generation. In *Neural Networks, 1995. Proceedings., IEEE International Conference on*, volume 2, pages 686–691. IEEE.

Hinton, G. E. (2012). A practical guide to training restricted boltzmann machines. In *Neural Networks: Tricks of the Trade*, pages 599–619. Springer.

Huang, Y., Akbari, H., Taha, H., and Rosenfeld, A. H. (1987). The potential of vegetation in reducing summer cooling loads in residential buildings. *Journal of climate and Applied Meteorology*, 26(9):1103–1116.

IEA. (2017). *World Energy Outlook 2017*. Organisation for Economic Co-operation and Development, OECD.

Jeong, H., Kim, H., Lee, S., and Dornfeld, D. (2006). Multi-sensor monitoring system in chemical mechanical planarization (cmp) for correlations with process issues. *CIRP Annals-Manufacturing Technology*, 55(1):325–328.

Jiang, Z., Liu, C., Akintayo, A., Henze, G. P., and Sarkar, S. (2017). Energy prediction using spatiotemporal pattern networks. *Applied Energy*, 206:1022–1039.

Jiang, Z. and Sarkar, S. (2015). Understanding wind turbine turbine interactions using spatiotemporal pattern network. In *Proceedings of ASME Dynamics Systems and Control Conference*.

Jin, S., Hou, Z., Chi, R., and Hao, J. (2014). A data-driven control design approach for freeway traffic ramp metering with virtual reference feedback tuning. *Mathematical Problems in Engineering*, 2014.

Kalvelage, K., Passe, U., Krejci, C., and Dorneich, M. C. (2015). Incorporation of future building operating conditions into the modeling of building–microclimate interaction: A feasibility approach. In *Sustainable Human–Building Ecosystems*, pages 150–158.

Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *CoRR*, abs/1412.6980.

Klein, L. A., Mills, M. K., and Gibson, D. R. (2006). Traffic detector handbook: -volume ii. Technical report.

Kuo, Y.-H. and Szeto, W. (2018). Smart transportation and analytics.

Lee, I., Sokolsky, O., Chen, S., Hatcliff, J., Jee, E., Kim, B., King, A., Mullen-Fortino, M., Park, S., Roederer, A., et al. (2012). Challenges and research directions in medical cyber–physical systems. *Proceedings of the IEEE*, 100(1):75–90.

Liu, C., Akintayo, A., Jiang, Z., Henze, G. P., and Sarkar, S. (2018). Multivariate exploration of non-intrusive load monitoring via spatiotemporal pattern network. *Applied Energy*, 211:1106–1122.

Liu, C., Ghosal, S., Jiang, Z., and Sarkar, S. (2016a). An unsupervised spatiotemporal graphical modeling approach to anomaly detection in distributed cps. In *Cyber-Physical Systems (ICCPS), 2016 ACM/IEEE 7th International Conference on*, pages 1–10. IEEE.

Liu, C., Ghosal, S., Jiang, Z., and Sarkar, S. (2017). An unsupervised anomaly detection approach using energy-based spatiotemporal graphical modeling. *Cyber-Physical Systems*.

Liu, C., Huang, B., Zhao, M., Sarkar, S., Vaidya, U., and Sharma, A. (December 2016b). Data driven exploration of traffic network system dynamics using high resolution probe data. In *Proceedings of IEEE Conference on Decision and Control, Las Vegas, NV*.

Lore, K. G., Stoecklein, D., Davies, M., Ganapathysubramanian, B., and Sarkar, S. (2015). Hierarchical feature extraction for efficient design of microfluidic flow patterns. In *Feature Extraction: Modern Questions and Challenges*, pages 213–225.

Luo, Z.-P., Berglund, L. J., and An, K.-N. (1998). Validation of f-scan pressure sensor system: a technical note. *Journal of rehabilitation research and development*, 35(2):186.

Meek, D. and Hatfield, J. (1994). Data quality checking for single station meteorological databases. *Agricultural and Forest Meteorology*, 69(1-2):85–109.

Minge, E. D., Peterson, S., Weinblatt, H., Coifman, B., and Hoekman, E. (2012). Loop-and length-based vehicle classification: Federal highway administration-pooled fund program [tpf-5 (192)]. Technical report, Minnesota Department of Transportation, Research Services.

Mukherjee, K. and Ray, A. (2014). State splitting and merging in probabilistic finite state automata for signal representation and analysis. *Signal processing*, 104:105–119.

Najafi, M., Gulp, C., and Langari, R. (2004). Enhanced auto-associative neural networks for sensor diagnostics (e-aann). In *Fuzzy Systems, 2004. Proceedings. 2004 IEEE International Conference on*, volume 1, pages 453–456. IEEE.

Pasqualetti, F., Franchi, A., and Bullo, F. (2018). Rigidity maintenance control for multi-robot systems. *Robotics: Science and Systems VIII*, 28:592–606.

Qu, F., Wang, F.-Y., and Yang, L. (2010). Intelligent transportation spaces: vehicles, traffic, communications, and beyond. *IEEE Communications Magazine*, 48(11).

Rao, C., Ray, A., Sarkar, S., and Yasar, M. (2009). Review and comparative evaluation of symbolic dynamic filtering for detection of anomaly patterns. *Signal, Image and Video Processing*, 3(2):101–114.

Ray, A. (2004). Symbolic dynamic analysis of complex systems for anomaly detection. *Signal Processing*, 84(7):1115–1130.

Sallans, B., Bruckner, D., and Russ, G. (2005). Statistical model-based sensor diagnostics for automation systems. *IFAC Proceedings Volumes*, 38(2):239–246.

Sarkar, S., Mukherjee, K., Sarkar, S., and Ray, A. (2013a). Symbolic dynamic analysis of transient time series for fault detection in gas turbine engines. *Journal of Dynamic Systems, Measurement, and Control*, 135(1):014506.

Sarkar, S., Sarkar, S., Virani, N., Ray, A., and Yasar, M. (2014). Sensor fusion for fault detection and classification in distributed physical processes. *Frontiers in Robotics and AI*, 1:16.

Sarkar, S. and Srivastav, A. (2016). A composite discretization scheme for symbolic identification of complex systems. *Signal Processing*, 125:156–170.

Sarkar, S., Srivastav, A., and Shashanka, M. (2013b). Maximally bijective discretization for data-driven modeling of complex systems. In *American Control Conference (ACC), 2013*, pages 2674–2679. IEEE.

Scofield, J. H. and Oberlin, O. (2009). A re-examination of the nbi leed building energy consumption study. In *International energy program evaluation conference, Portland, OR*.

Solo, V. (2008). On causality and mutual information. In *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, pages 4939–4944. IEEE.

Song, M., Tarn, T.-J., and Xi, N. (2000). Integration of task scheduling, action planning, and control in robotic manufacturing systems. *Proceedings of the IEEE*, 88(7):1097–1107.

Sun, Y., Heo, Y., Tan, M., Xie, H., Jeff Wu, C., and Augenbroe, G. (2014). Uncertainty quantification of microclimate variables in building energy models. *Journal of Building Performance Simulation*, 7(1):17–32.

Swan, L. G. and Ugursal, V. I. (2009). Modeling of end-use energy consumption in the residential sector: A review of modeling techniques. *Renewable and sustainable energy reviews*, 13(8):1819–1835.

Tamilselvan, P., Wang, P., and Youn, B. D. (2011). Multi-sensor health diagnosis using deep belief network based state classification. In *ASME 2011 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, pages 749–758. American Society of Mechanical Engineers.

Thonhofer, E., Palau, T., Kuhn, A., Jakubek, S., and Kozek, M. (2018). Macroscopic traffic model for large scale urban traffic network design. *Simulation Modelling Practice and Theory*, 80:32–49.

Vico, M., Zibman, S., and Stefanovic, A. (2015). Vehicular camera and method for periodic calibration of vehicular camera. US Patent 9,150,155.

Wang, R., Zhang, L., Sun, R., Gong, J., and Cui, L. (2011). Easitia: A pervasive traffic information acquisition system based on wireless sensor networks. *IEEE Transactions on Intelligent Transportation Systems*, 12(2):615–621.

Wang, S. and Chen, Y. (2004). Sensor validation and reconstruction for building central chilling systems based on principal component analysis. *Energy Conversion and management*, 45(5):673–695.

Wang, Z., Wang, Y., and Srinivasan, R. S. (2018). A novel ensemble learning approach to support building energy use prediction. *Energy and Buildings*, 159:109–122.

Wells, T. J., Smaglik, E. J., and Bullock, D. M. (2008). Health monitoring procedures for freeway traffic sensors, volume 1: Research report. *Joint Transportation Research Program*, page 318.

Wenjie, C., Lifeng, C., Zhanglong, C., and Shiliang, T. (2005). A realtime dynamic traffic control system based on wireless sensor network. In *Parallel Processing, 2005. ICPP 2005 Workshops. International Conference Workshops on*, pages 258–264. IEEE.

Wibral, M., Rahm, B., Rieder, M., Lindner, M., Vicente, R., and Kaiser, J. (2011). Transfer entropy in magnetoencephalographic data: Quantifying information flow in cortical and cerebellar networks. *Progress in biophysics and molecular biology*, 105(1):80–97.

Wu, J.-q., Ang, J.-C. F., and Roitberg, L. J. J. (2018a). System and method of directional sensor calibration. US Patent 9,863,785.

Wu, L., Liu, C., Huang, T., Sharma, A., and Sarkar, S. (2018b). Traffic sensor health monitoring using spatiotemporal graphical modeling. *IJPHM*.

Xiao, F., Wang, S., and Zhang, J. (2006). A diagnostic tool for online sensor health monitoring in air-conditioning systems. *Automation in Construction*, 15(4):489–503.

Yang, H.-T., Huang, C.-M., and Huang, C.-L. (1996). Identification of armax model for short term load forecasting: an evolutionary programming approach. *IEEE Transactions on Power Systems*, 11(1):403–408.

Yang, X., Zhao, L., Bruse, M., and Meng, Q. (2012). An integrated simulation method for building energy performance assessment in urban environments. *Energy and Buildings*, 54:243–251.

Zhang, J., Tai, L., Xiong, Y., Liu, M., Boedecker, J., and Burgard, W. (2018). Vr goggles for robots: Real-to-sim domain adaptation for visual control. *arXiv preprint arXiv:1802.00265*.