# Traffic System Anomaly Detection using Spatiotemporal Pattern Networks

Tingting Huang[1], Chao Liu[2], Anuj Sharma[1], and Soumik Sarkar[2]

[1]*Department of Civil, Construction and Environmental Engineering, Iowa State University, Ames, Iowa, 50010, United States*
*thuang1@iastate.edu*
*anujs@iastate.edu*

[2] *Department of Mechanical Engineering, Iowa State University, Ames, Iowa, 50010, United States*
*cliu5@tsinghua.edu.cn*
*soumiks@iastate.edu*

**ABSTRACT**

Traffic dynamics in the urban interstate system are critical in terms of highway safety and mobility. This paper proposes a systematic data mining technique to detect traffic system-level anomalies in a batch-processing fashion. Built on the concepts of symbolic dynamics, a spatiotemporal pattern network (STPN) architecture is developed to capture the system characteristics. This novel spatiotemporal graphical modeling approach is shown to be able to extract salient time series features and discover spatial and temporal patterns for a traffic system. An information-theoretic metric is used to quantify the causal relationships between sub-systems. By comparing the structural similarity of the information-theoretic metrics of the STPNs learnt from each day, a day with anomalous system characteristics can be identified. A case study is conducted on an urban interstate in Iowa, USA, with 11 roadside radar sensors collecting 20-second resolution speed and volume data. After applying the proposed methods on one-month data (Feb. 2017), several system-level anomalies are detected. The potential causes that include inclement weather condition and non-recurring congestion are also verified to demonstrate the efficacies of the proposed technique. Compared to the traditional predefined performance measures for the traffic systems, the proposed framework has advantages in capturing spatiotemporal features in a fast and scalable manner.

## 1. INTRODUCTION

Traffic systems are complex, interactive and dynamic. Both temporal and spatial relationships that exist among multiple attributes and different sub-systems in a traffic system need to be extracted for effective performance monitoring. From a traffic operation perspective, establishing a reliable and intelligent transportation system could benefit both system planners and users, who relies highly on data. However, as a result of rapidly growing data, how to efficiently mine the hidden pattern of those data and further monitoring the health of the system becomes important.

In transportation research, many studies have been done in detecting incidents. Margreiter (2016) used Bluetooth reidentification techniques to estimate travel time and further detected congestion/incident by a thresholding method. The authors used 80 km/h as speed threshold for warning and combined both number of warnings and 60 km/h speed threshold to detect incidents. Besides the simple fixed thresholding method, some other statistical method was also employed. Chakraborty, Hess, Sharma and Knickerbocker (2017) used an outlier-based method to explore more from historical data then set up a dynamic threshold of speed for detection. Other than threshold-based method, Tang and Gao (2005) proposed a combined method of the nonparametric regression and standard deviation algorithm to detect incidents and tested it in simulation. Jin and Ran (2009) utilized the fundamental diagrams in traffic flow theory to identify the freeway incidents, and improved it by introducing uncongested and congested regime shifts in the diagrams.

As artificial intelligence was applied widely in recent decades, there have been also many machine learning methods applied in traffic incident detection. Many techniques like decision tree, support vector machine (SVM) and neural network were practiced. Chen and Wang (2009) used traffic volume, speed, vehicle headway and sensor occupancy data to implement decision tree learning and tested it in a simulated environment. Regarding SVM,

Yuan and Cheu (2003) used two different non-linear kernel SVMs to train and test in simulated incidents data. To optimize the parameters for SVM, Yao, Hu, Zhang and Jin (2014) employed the tabu search algorithm to achieve more accurate classification. Moreover, Li, He, Zhang and Yang (2016) proposed a bagging SVM for classifying highway incidents. They bootstrapped several subsets to train SVMs, then used majority voting to ensemble them. Another research done by Kim and Wang (2016) used Bayesian networks to detect and predict highway congestion. Besides the traffic flow characteristics like speed and volume of the flow, they also used weather condition and time of day as inputs.

There are also many studies utilizing neural network to identify the incidents. Ritchie and Cheu (1993) used traffic data from simulation and train a multi-layer neural network to detect freeway incidents. To improve the detection performance, Abdulhai and Ritchie (1999) then applied a modified form of Bayesian-based neural network and achieved faster training and higher performance than previous architecture. Further, Adeli and Karim (2000) proposed a fuzzy-wavelet radial basis function neural network to classify the incidents, it also achieved high detection rate and low false alarms in both real world and simulated data.

However, these previous machine learning methods adopted in transportation area tend to be supervised learning, which requires expensive labeled data and more variables to train the model. Moreover, the common objective of these research is still trying to detect isolated incident at traffic operation level, which is finding the location and time of an incident. In terms of system-wide anomaly, they might ignore other factors resulting in traffic pattern changes, such as adverse weather condition.

This work aims to use an unsupervised learning method to detect anomalies from a system-wide perspective. The motivation of system-wide anomaly detection is that an event occurrence may not always lead to a severe impact on system. Thus, it is important to build a health monitoring process that focuses on the system dynamics, in this case, the traffic flow dynamics. The approach in this work is intended to capture system-wide anomalies, other than the events that only affect the local dynamics, and this kind of method is more robust with noise and disturbances in the system.

To achieve an unsupervised, systematic learning, we apply a novel data-driven method based on spatiotemporal pattern network (STPN). This framework has been successfully applied in solving different real-world engineering problems. For example, STPN has been used for bridge damage detection in structural health monitoring (Liu, Gong, Laflamme, Phares, & Sarkar, 2017). Researchers proposed an approach based on STPN to extract patterns from dense sensor network, and applied it on damage detection in a small bridge network. Results showed that the approach could capture the spatiotemporal features, localize the damage and it can be implemented in real-time. Another application of STPN framework is wind turbine power prediction (Jiang, Liu, Akintayo, Henze, & Sarkar, 2017). Researchers used STPN models to extract spatiotemporal features and capture causal dependencies. They also predicted the power for one wind turbine based on the observation from another wind turbine and achieved a high degree of accuracy. Moreover, one research (Liu, Huang, Zhao, Sarkar, Vaidya, & Sharma, 2016) has been done using STPN to explore traffic dynamics on an interstate, which demonstrates a good application of STPN in traffic system.

**Contributions** This study applies a novel framework, the spatiotemporal pattern network, to detect the traffic system anomaly. In contrast with the traditional transportation research methods, it captures the spatiotemporal features of traffic flow and discovers the causal relationships between the sub-systems. Also, it only learns from data instead of using traditional predefined measures, which helps mitigate the impacts from arbitrary rules. Besides, compared to the machine learning methods used previously, it is also fast and easy to implement without the need of expensive labeled data. In addition, it does not involve much site-specific information, which makes it more scalable.

In this study, we used the high-resolution, 2-dimensional real historical traffic data over one month from 11 roadside radar sensors on Interstate 35/80 in Des Moines, Iowa. The proposed graphical modeling approach is used to extract the pattern of traffic dynamics and detect the anomalies. Several anomalies are identified and potential practical causes are also investigated in the case study.

This work could also be extended into an online detection application. Some related work has already been performed by Lin, Liu, Huang, Sarkar and Sharma (2017). Although an online detection is very useful as sending early warnings to road users, there is also a need of extracting long term trend by using batch processing focused on historical data. It is critical to decision-makers examining the different impacts from past events and preparing appropriate reaction plan accordingly.

This paper has 6 sections including introduction. Section 2 introduces the framework of STPN and the metrics for STPN; Section 3 focuses on the problem formulation, including data description and STPN learning. Section 4 discusses the results from STPN evaluation and anomaly detection. Section 5 demonstrates some additional works including application on original data and scalability test. Section 6 concludes this paper along with future research directions.

## 2. METHODOLOGY

### 2.1. Spatiotemporal Pattern Network (STPN)

Built on the concepts of Symbolic Dynamics Filtering, a spatiotemporal feature extraction scheme, STPN, is constructed to discover and represent sub-system behavior and causal interactions among the sub-systems (Sarkar, Sarkar, Virani, Ray, & Yasar, 2014; Jiang & Sarkar, 2015; Liu, Ghosal, Jiang, & Sarkar, 2017). The fundamental concept of STPN, symbolic dynamic filtering, has advantages in extracting features from time series data (Rao, Ray, Sarkar, & Yasar, 2009). It is able to use symbol sequence to approximate a $D$-Markov machine to capture the features in the process.

Data abstraction (discretization and symbolization) is the first step to create discrete symbol sequences from continuous data. Thus, the system is analyzed in the symbolic space instead of the continuous space. The discretization and symbolization of time series data is done by partitioning. The general idea of partitioning is, for a given time series data $T$ with n samples, transform $T$ into symbol sequence $S$ with $k$ partitions where $k \leq n$. There are several partitioning algorithms could be used, such as uniform partitioning (UP), maximum entropy partitioning (MEP), maximum migration partitioning (MMP), symbolic false nearest neighbor partitioning (SFNNP), etc. (Jin, Sarkar, Mukherjee, & Ray, 2009; Sarkar, Srivastav, & Shashanka, 2013; Sarkar & Srivastav, 2016). In this study, since traffic system is closely related to the physical world, to reflect the relationship between traffic data and public knowledge, a customized UP was proposed to transform all the time series into symbol sequences with 6 partitions. The details will be elaborated in case study.

Another assumption in this modeling approach is that we can approximate a symbol sequence as a Markov chain of order $D$. Thus, a $D$-Markov machine (or $xD$-Markov machine for multivariate time series) could be built to analyze the temporal features ($xD$-Markov machine is for extracting spatial features).

A $D$-Markov machine is a probabilistic finite state automata (PFSA) using finite history of $D$ symbols as one state. It is formally defined as follows (Sarkar et al., 2014).

- $D$ is the depth of the Markov machine;
- $Q$ is the finite set of states with cardinality $|Q| \leq |\Sigma|^D$, the states are represented by equivalence classes of symbol strings of maximum length $D$ where each symbol belongs to alphabet $\Sigma$;
- and $\delta: Q \times \Sigma \to Q$ is the state transition function that satisfies the condition that if $|Q| = |\Sigma|^D$, there exist $\alpha, \beta \in \Sigma$ and $x \in \Sigma^\star$ such that $\delta(\alpha x, \beta) = x\beta$ and $\alpha x, x\beta \in Q$.

where $Q$ is a non-empty finite set with cardinality $|Q| \leq \infty$, called set of states; $\Sigma$ is a non-empty finite set with cardinality $|\Sigma| \leq \infty$, called symbol alphabet; and $\Sigma^\star$ is the collection of all finite-length strings with symbols from $\Sigma$.

As defined above, a $D$-Markov machine estimates the probability of occurrence of a new symbol given the last $D$ symbols for one symbol sequence, thus, it can capture the causal effects of one symbol sequence on another symbol sequence (Jiang & Sarkar, 2015).

To determine the cross-dependence, an $xD$-Markov machine is defined as follows (Sarkar et al., 2014).

Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be the PFSAs corresponding to symbol sequence $\{s_1\}$ and $\{s_2\}$ respectively. An $xD$-Markov machine is defined as a 5-tuple $\mathcal{M}_{1\to2} \triangleq (Q_1, \Sigma_1, \Sigma_2, \delta_1, \widetilde{\Pi}_{12})$ such that:

- $Q_1 = \{q_1, \dots, q_{|Q_1|}\}$ is the state set of symbol sequence $\{s_1\}$;
- $\Sigma_1 = \{\sigma_0, \dots, \sigma_{|\Sigma_1|-1}\}$ and $\Sigma_2 = \{\sigma_0, \dots, \sigma_{|\Sigma_2|-1}\}$ are the alphabet sets of symbol sequence $\{s_1\}$ and $\{s_2\}$ respectively;
- $\delta_1: Q_1 \times \Sigma_1 \to Q_1$ is the state transition function that maps the transition in symbol sequence $\{s_1\}$;
- $\widetilde{\Pi}_{12}$ is the symbol generation matrix of size $Q_1 \times \Sigma_2$; the $ij^{th}$ element of $\widetilde{\Pi}_{12}$ denotes the probability of finding the symbol $\sigma_j$ in $\{s_2\}$ while making a transition from the state $q_i$ in $\{s_1\}$.

With this setup, STPN is defined as a 4-tuple $W_D$:

$$W_D \equiv (Q^A, Q^B, \Pi^{AB}, \Lambda^{AB}) \tag{1}$$

such that:

- $A$ and $B$ are representing two sub-systems (nodes) of STPN;
- $Q^A$ and $Q^B$ are the state set correspondingly;
- $\Pi^{AB}$ indicates the transition matrix from $A$ to $B$;
- and $\Lambda^{AB}$ is a metric for quantifying the relational pattern from $A$ to $B$.

Figure 1 demonstrates the structure of STPN model. In Fig. 1, $\Pi^{AA}$ and $\Pi^{BB}$ are the transition matrices representing the self-relations for system $A$ and system $B$ correspondingly, which are also referred to atomic patterns (APs). While $\Pi^{AB}$ and $\Pi^{BA}$ are the transition metrics reflecting cross relations from $A$ to $B$ and from $B$ to $A$, which are called relational patterns (RPs). Formally the transition matrix is derived by:

$$\pi_{\alpha\beta}^{AB} \coloneqq P(S_{i+1}^B = \beta \,|\, S_i^A = \alpha) \; \forall i \tag{2}$$

where $\alpha \in Q^A$ and $\beta \in Q^B$; $\pi_{\alpha\beta}^{AB}$ is the probability of transiting from state $\alpha$ in system $A$ to state $\beta$ in system $B$.

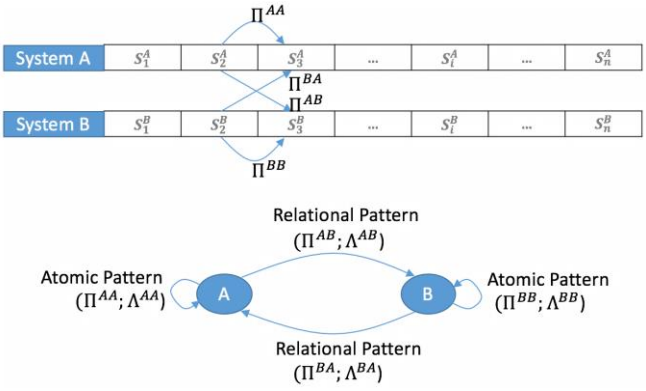



Figure 1: Extraction of atomic patterns and relational patterns of STPN

The APs intend to extract the state transitions in a sub-system itself, and the RPs describe the state transition from a sub-system to another. Using Eq. (2), the transition probabilities can be computed and represent the patterns (APs and RPs).

To quantify the APs and RPs in STPN, $\Lambda^{AB}$ is defined. Here, an information theoretic metric could be used (Solo, 2008; Wibral, Rahm, Rieder, Lindner, Vicente, & Kaiser, 2011). There are several metrics available, such as transfer entropy and mutual information. In this study, the mutual information (MI) is used.

### 2.2. Mutual Information based Metric

In this study, we define the MI for APs and RPs as follows (RP from system $A$ to $B$ is used as instance).

$$I^{AB} = H(S_{i+1}^B) - H(S_{i+1}^B \mid S_i^A) \tag{3}$$

where

$$H(S_{i+1}^B) = -\sum_{\beta}^{Q^B} P(S_{i+1}^B = \beta)\, log_2 P(S_{i+1}^B = \beta)$$

$$H(S_{i+1}^B \mid S_i^A) = -\sum_{\alpha}^{Q^A} P(S_i^A = \alpha)\, H(S_{i+1}^B \mid S_i^A = \alpha)$$

$$H(S_{i+1}^B \mid S_i^A = \alpha) = -\sum_{\beta}^{Q^B} P(S_{i+1}^B = \beta \mid S_i^A = \alpha) \cdot log_2 P(S_{i+1}^B = \beta \mid S_i^A = \alpha)$$

This MI based metric is used to measure the capability of predicting the dynamics of one sub-system from past observations of another sub-system dynamics or itself.

### 2.3. Structural Similarity

In this study, we treat each sensor on the road as one node or sub-system of STPN. Thus, an $N \times N$ MI-matrix ($N$ is number of sensors) could be obtained to represent the patterns in STPN. As we examine the data in a daily basis, we would obtain $M$ MI-matrices in total during study time period (here $M = 28$), and a comparison method is needed. Here we adopt an index called structural similarity (SSIM) from image processing. SSIM (Wang, Bovik, Sheikh, & Simoncelli, 2004) is focusing on the structural information of an image, like the pixels have strong inter-dependencies especially when they are spatially close. Formally it is defined as follows (Wang et al., 2004).

$$S(x,y) = \left(\frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}\right)^{\alpha} \left(\frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}\right)^{\beta} \left(\frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}\right)^{\gamma} \tag{4}$$

where

- $\mu_x$ and $\mu_y$ are the mean of $x$ and $y$ respectively;
- $\sigma_x^2$ and $\sigma_y^2$ are the variance of $x$ and $y$ respectively;
- $\sigma_{xy}$ is the covariance of $x$ and $y$;
- $C_1$, $C_2$, and $C_3$ are used to stabilize the division if denominator is near 0;
- $C_1 = (k_1 L)^2$, $C_2 = (k_2 L)^2$ and $C_3 = C_2/2$ with $k_1, k_2$ and $L$ being constant;
- $\alpha$, $\beta$ and $\gamma$ are weights for combining those comparative measures with $\alpha, \beta, \gamma > 0$.

SSIM measures the local quality/distortion between two images using a sliding window and combines the results to a single value as the index of one image's quality related to another image (Wang et al., 2004). Although the SSIM index is designed for comparing images, it has been shown to be useful in computing the similarity of features (Liu, Jiang, & Yang, 2014). For our $N \times N$ MI-matrix, which could be treated as images, the SSIM index is efficient in terms of feature extraction and comparison. Here, SSIM index is not related to a specific traffic condition. It is used as a metric to compare the similarity of features (represented by MI matrix for each day), where a low SSIM index indicates the traffic conditions represented by the MI matrices are different.

## 3. PROBLEM FORMULATION

In this study, we utilized real word traffic data from sensors, and applied STPN for anomaly detection. Figure 2 depicts the basic work flow.

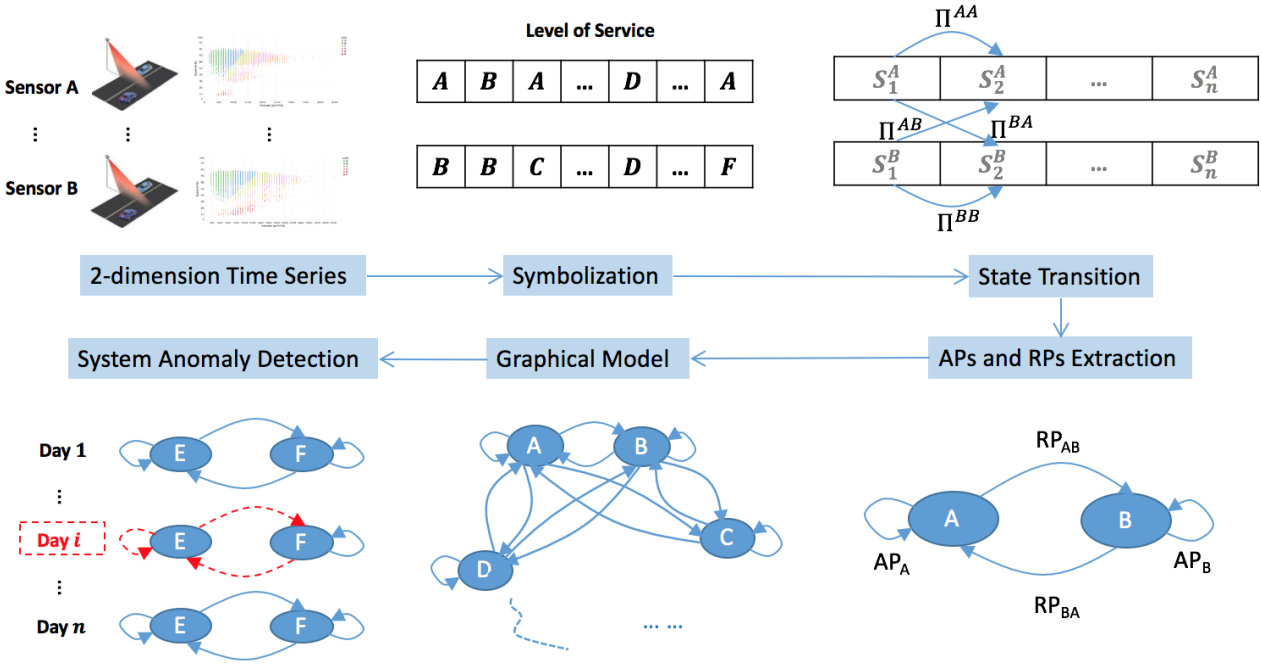



Figure 2. Construction and learning of STPNs for anomaly detection from daily traffic data

As shown in Fig. 2, the multivariate time-series data collected from the sensors are first partitioned into symbols and then state sequences are generated. The state transition matrices are then obtained using $D$-Markov machine ($xD$-Markov machine). The patterns are then evaluated using information based metric (mutual information in this work) and daily graphical models are formed. The system-wide anomaly affects the patterns ("Day $i$" marked at the bottom-left panel) and can be detected through comparing the changes of the mutual information metrics.

### 3.1. Data Preparation

This study used traffic data collected from 11 radar sensors on I-35/80 WB through Des Moines urban area (speed limit is unchanged segment to segment). The location of each sensor is shown in Fig. 3.

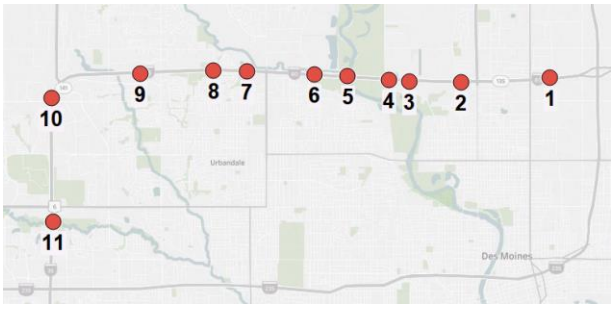



Figure 3. Location of studied sensors on I-35/80 westbound, labeled as order in traveling direction

These sensors are labeled by their orders in terms of traveling direction. Speed and volume data in 20-second intervals were obtained from these sensors. In this case study, we took February 2017 as the study time period.

As the model requires continuity in time series data, we need to preprocess the data when there was no vehicle present. Since this situation happened at night at most times, thus, we excluded night time (11pm-5am) data from the daily data set. For any other missing values in some sensor, we linearly interpolated the value by using the speed and volume at closest timestamps before and after. However, if a start or end value is missed, the interpolation will fail. Thus, we also used the smallest overlapping time period in each day with all the sensors available. After the data preprocessing, this system has two-dimensional time series data with 11 nodes for 28 days.

### 3.2. Symbolization

This study uses custom domain knowledge based partitioning to transform the continuous time-series data into symbol sequence. In Highway Capacity Manual (HCM) (Transportation Research Board, 2000), level of service (LOS) is a quality measure regarding operational conditions under different traffic flows.

There are 6 lettered LOS from "A" to "F", with "A" representing the best and "F" the worst. Different types of road facilities require different methods to compute LOS. In this study, we employ the method for freeway LOS

calculation based on traffic density. The traffic density is defined by the number of passenger cars presenting in one kilometer one lane. The computation of density follows:

$$D = V/S \tag{5}$$

where V is the flow rate (in pc/hr/ln) and S is the average speed (in km/hr).

The LOS is determined by the density value. Table 1 lists the LOS criteria for basic freeway segments from HCM.

This LOS-based custom partitioning algorithm is applied on the entire dataset, and the result are illustrated in Fig. 4. After symbolization, the continuous multivariate time series data are discretized into univariate 6-symbol sequences.

| LOS | Density (pc/km/ln) |
|---|---|
| A | [0, 6.83] |
| B | (6.83, 11.18] |
| C | (11.18, 16.15] |
| D | (16.15, 21.74] |
| E | (21.74, 27.95] |
| F | (27.95, maximum] |

Table 1. Freeway LOS criteria

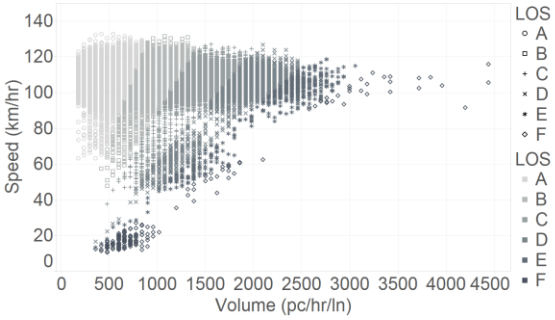



Figure 4. Traffic data partitioning via LOS rules

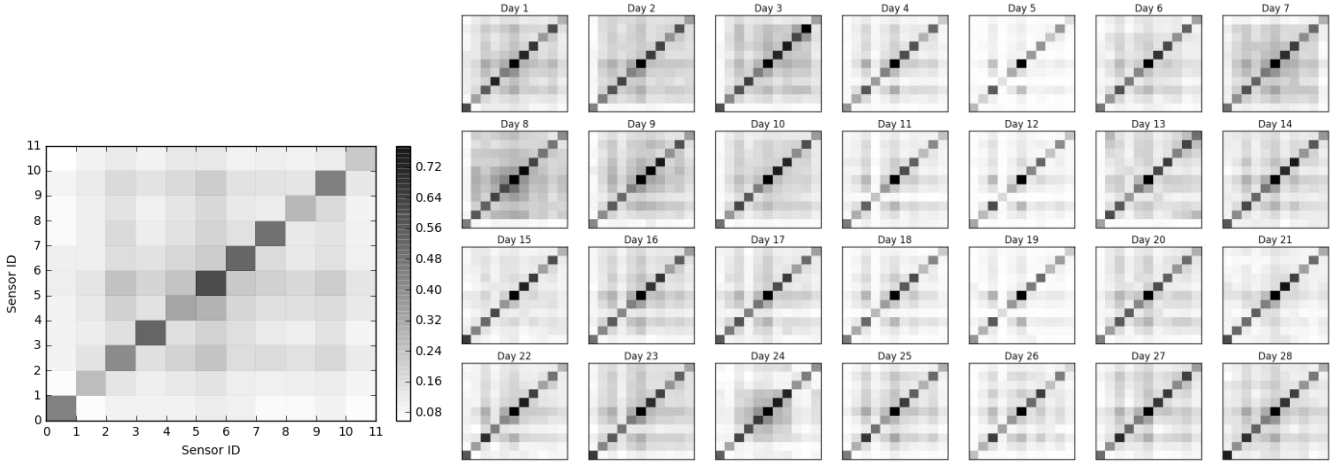
### 3.3. MI Calculation and STPN Evaluation

After getting the symbol sequences from each sensor, we treated them as Markov chains of order $D$ ($D$=1 in this work), and computed the 1-step transition matrices, in order to form the STPN with less complex computation. Further, to quantify the connectivity among those sub-systems (i.e. sensors in this case), MI was calculated on those transition matrices by using Eq. (3). An example of MI results is shown in Fig. 5. The Fig. 5 (a) is just showing the quantification of Day 1's STPN, in which the darker color represents higher MI between sensors. And Fig. 5 (b) is showing all the MI matrices in study period with the same color scheme in Fig. 5 (a).

The higher value of MI from a to b indicates the more information obtained in sensor b is through sensor a. In other words, MI represents how well one sensor could predict another. Together they formed the whole metrics of a pattern network, which could reflect the system dynamics.

To efficiently compare those MI-matrices on STPNs, the SSIM index is calculated using default window size 7 and uniform filter. SSIM is symmetric, which means the SSIM for Day 1 to Day 2 is the same as for Day 2 to Day 1. Since the comparison strategy is sensitive to the baseline selection, in this study, we use the following comparison strategy: for a certain day, calculate all the SSIM indices from this day to the other days, then use the average value as the index for it.

To identify the anomalous days, here we use 85% of the maximum SSIM value as the threshold rather than a percentile thresholding for anomalies. The reason for setting this threshold includes: (i) the SSIM on any anomalous days should be away from the best condition (maximum SSIM); (ii) we should avoid using percentile, which will maintain a fixed portion of days in every month to be anomalies. The results are illustrated in Fig. 6.




a) Day 1 example with sensor ID b) All days' presentation

Figure 5. Information based metrics, each small block represents the MI between that pair of sensors

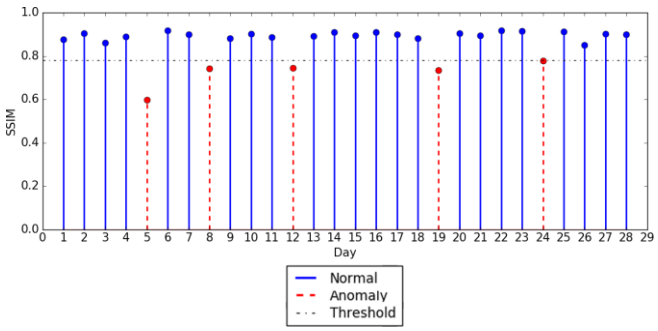



Figure 6. Average SSIM from STPN in each day, dashed ones are identified as anomalous days

## 4. RESULTS DISCUSSION

As shown in Fig. 6, Feb. 5th had a significant drop in average similarity to others. Other days like Feb. 8th, Feb. 12th, Feb. 19th and Feb. 24th also had less similarities. Motivated by the potential day-of-week seasonality (low SSIM on Feb. 5th, 12th, 19th) and a prior knowledge of traffic variation in terms of day of week (especially weekday vs. weekend), we further explore the patterns by comparing them at the day of week level.

Figure 7 shows the average SSIM for each day in day of week level. For example, Wednesday in Week 1 (Feb. 1st) obtained its SSIM index by averaging SSIM indices comparing with all other Wednesdays. Thus, as Fig. 7 indicates, Wednesdays in the study period show relatively low and diverse SSIM values, and Saturdays have a more stable pattern.

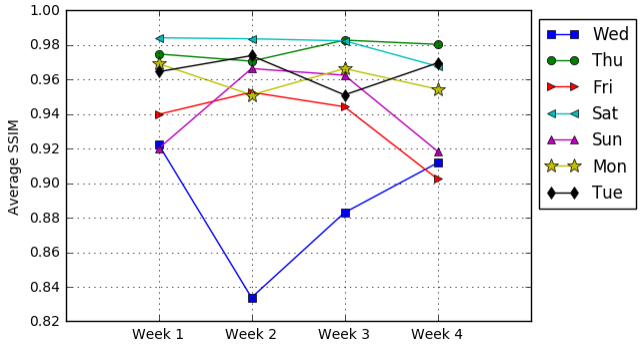



Figure 7. Average SSIM from STPN by day of week

To associate the patterns with the real-world situation, a heat map has been generated by using the interpolated data set. Figure 8 visualizes the LOS in the whole system every day, by using vertical axis to represent sensors and horizontal axis as time of day.

### 4.1. Events: Adverse Weather and Crash

From Fig. 8, it could be seen that on Feb. 8th (Wednesday, Week 2) and Feb. 24th (Friday, Week 4), there were unusually bad LOS present in morning and afternoon peak hours. By checking the historical weather information (Weather Underground, 2017), it shows that there were snowfall events in those two days. Thus, the inclement weather may cause the anomalous pattern in those days since it is reasonable to assume the motorists on highway could be affected by heavy snows.

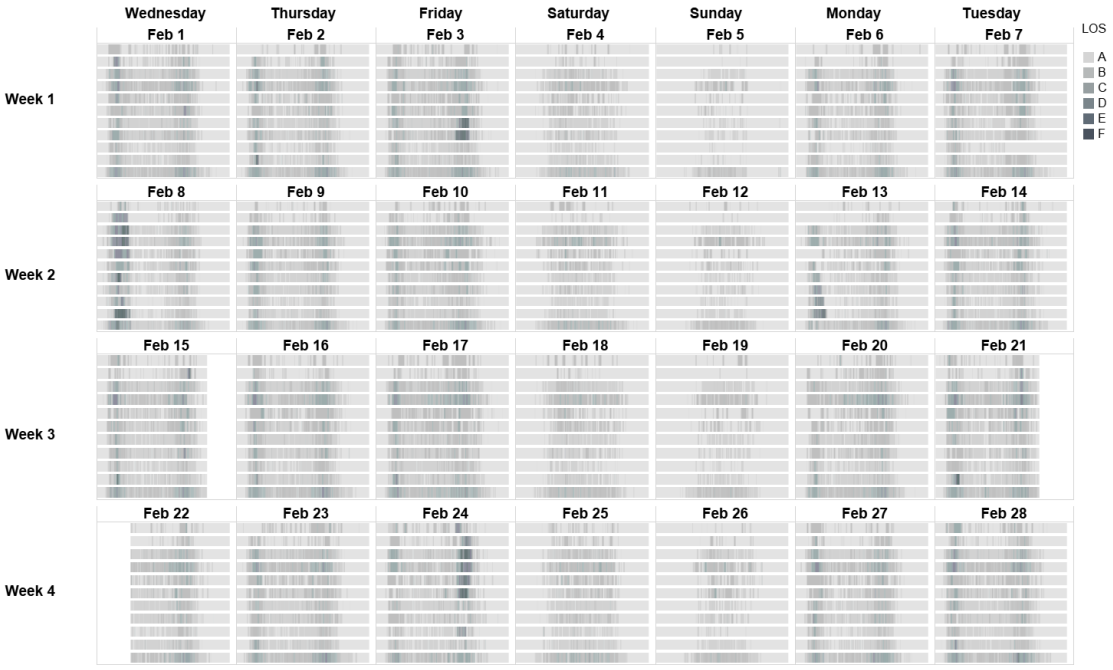



Figure 8. LOS heat map from the traffic system in each day, with x-axis represents time of day and y-axis represents sensors

Another data source that we have access to is the event reports from Iowa DOT Traffic Management Center. Table 2 shows the number of events (focused on crash only) on each day in study time period on I-35/80 WB. Here it also shows on Feb. 8th and Feb. 24th, there were 2 and 5 crashes respectively. Therefore, we find that multiple vehicle crashes may contribute in making the system anomalous in those days as well.

Although the weather information and event reports could help us to verify the system anomalies we detected, they could not replace STPN to detect system anomaly directly. The reason why they are not suitable is that bad weather or crashes do not always severely affect the traffic system. For example, in Table 2, we could see that on Feb. 25th there were 2 multiple vehicle crashes. However, it still has a relatively high similarity with other Saturdays shown in Fig. 7 and Fig. 8 (Saturday, Week 4). The reason could be less volume in the weekend. Note that it is also not identified as a system-level anomaly by the proposed STPN scheme. In this context, STPN shows advantages in detecting the system-wide anomaly for the traffic system with fewer false alarms (the false alarms that may be reported when deploying weather or event information).

Note that such system-level anomalies arise from a complex combination of multiple factors involving weather, traffic states and incidents that can be highly non-intuitive in nature. Therefore, a multivariate automated feature extraction scheme such as STPN can be more effective compared to a rule-based univariate scheme for real life deployment.

### 4.2. Anomaly in Weekends

As shown in Fig. 6, some Sundays (Feb. 5th, Feb. 12th and Feb. 19th) were identified as anomaly due to the low similarity with all other days. Although another Sunday (Feb. 26th) was not detected as anomaly, it had relatively low similarity as well. Associated with Fig. 8, it could be seen that there were no obvious peak hours occurred on Sundays comparing to other days. This kind of anomaly captured by STPN is caused by different traffic pattern at weekends. Thus, it is necessary to differentiate the anomalies STPN detected in weekends from weekdays due to the nature of traffic pattern change by day of week. It would be beneficial that conducting the health monitoring on weekday and weekend separately.

| **Event Type** | 2-1 | 2-3 | 2-6 | 2-7 | 2-8 | 2-13 | 2-21 | 2-24 | 2-25 |
|---|---|---|---|---|---|---|---|---|---|
| 1 Vehicle Crash | 1 | | | 1 | 1 | | | 2 | |
| 2 Vehicle Crash | | 1 | 1 | | 1 | | | 2 | 2 |
| 3+ Vehicle Crash | | 1 | | | | 1 | 1 | 1 | |

Table 2. Number of crashes by date from event reports

In addition, Sunday trend is not as stable as Saturday shown in Fig. 7. Because there are only 4 data points in each day of week, it is not easy to determine and finalize the trend, especially in low volume weekends. Thus, a long-term monitoring of weekend trend is necessary and will be considered in the future work.

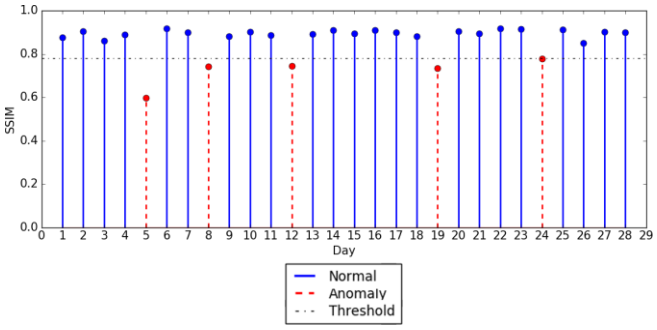
## 5. ADDITIONAL STUDY

### 5.1. Comparison with Original Information Similarity

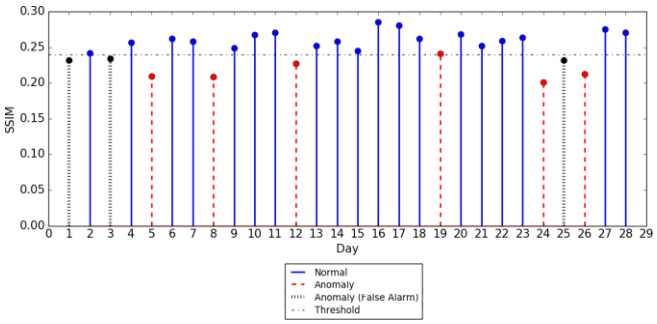
In addition, we also consider if simple image analysis of LOS heat maps (original information without STPN) over different days can be effective in anomaly detection. We compute the SSIM index directly based on the LOS heat maps (Fig. 8) and use the same averaging and thresholding strategy. The comparison with STPN results are shown in Fig. 9.

Compared to Fig. 9 (a), which is obtained from STPN, Fig. 9 (b) shows more fluctuations. Also, we observe that multiple nominal days and anomalous days are too close around the threshold, which indicates the results tend to be quite sensitive with the threshold. Also, using heat map directly may generate more false alarms.

Further investigation is also made regarding the distributions of SSIM under normal and anomalous conditions. Since the sample size is limited, here we assume




a) SSIM from STPN (same as Fig. 6)




b) SSIM from LOS heat map

Figure 9. Comparison of average SSIM from STPN and LOS. Dotted line in b) shows the additional false alarms

that the SSIM values follow Gaussian distributions just for illustration purpose. Here we also assume that the severe crash days and weekends have different characteristics than regular traffic flow. Thus, we could illustrate the SSIM distributions based on our benchmark from domain knowledge. Figure 10 shows the comparison of SSIM distributions from STPN and LOS heat map.

In Fig. 10 (a), STPN results show less variance in distribution under normal condition than anomalous condition and two distributions are well separated. Both of these characteristics are extremely useful for efficient anomaly detection with low false alarm. However, in LOS heat map results (Fig. 10 (b)), distributions under normal condition and anomalous conditions are not as well separated. This illustrates the need for a sophisticated scheme such as STPN for detecting traffic system-wide anomalies in a robust fashion.

### 5.2. Scalability Analysis

One additional case study was also conducted to test the scalability of this method. Data from the same corridor in January 2017 were used. By using the proposed methodology, Fig. 11 demonstrates both the SSIM from STPN results and the original LOS information.

By checking the weather information (Weather Underground, 2017), those anomaly days (in Fig. 11(a)) have low visibility with high perception, which impact the driver behaviors more significant than other days. Also, if we simply use the structural similarity method to extract information from original LOS, more variant SSIM values and more false alarms will be generated as shown in Fig.11(b). Thus, we still suggest to use proposed method to extract features and capture causal dependencies to conduct a robust detection.

This additional case implied that the proposed method could be easily implemented on other cases without rebuilding model to accommodate any site-specific or time-specific characteristics in transportation system.

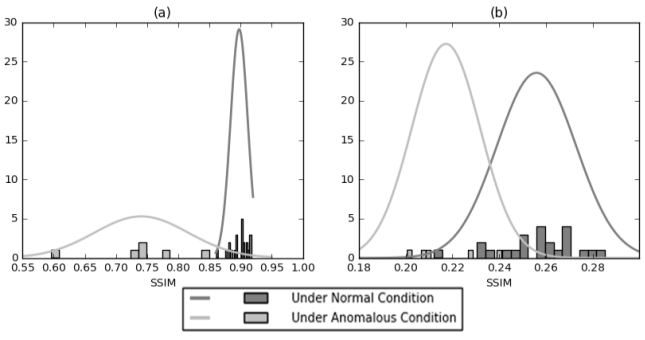
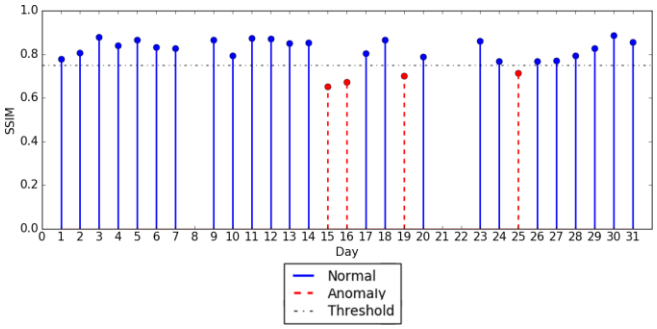
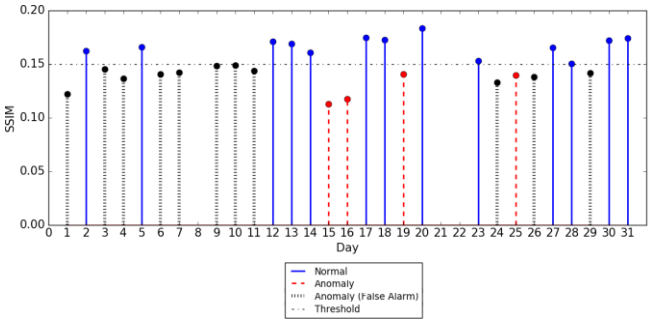



Figure 10. Comparison of SSIM distributions from STPN and LOS (a) SSIM from STPN (b) SSIM from LOS




a) SSIM from STPN (Additional case)




b) SSIM from LOS heat map (Additional case)

Figure 11. Additional case: comparison of average SSIM from STPN and LOS, blank space indicates missing data on that day

## 6. CONCLUSIONS AND FUTURE WORK

This research explored the traffic system dynamics and proposed a health monitoring approach. Built on concepts of symbolic dynamics, a spatiotemporal pattern network framework was presented to capture the system dynamics, and a mutual information based metric was used to quantify the causal relationship (atomic pattern and relational pattern) between sensors in the system. To compare the similarity of the information based metrics of the STPNs and further detect the anomaly, an SSIM measure was adopted to measure the similarity. Based on the assumption that the system-wide anomalies lead to significant variation in the patterns of the STPNs, the less similar patterns were identified as system anomaly.

This study applied the proposed method on one-month traffic data collected from 11 roadside radar sensors along I-35/80 WB in Iowa. By constructing STPN on daily traffic data, and comparing them in day of week level, several system anomalies with low similarities were detected. Associating weather and incident information, the potential causes of those system were also verified. It shows that the inclement weather and crashes could impact the system dynamics but not necessarily.

This paper employs and customizes the probabilistic graphical modeling method to solve a traffic system

problem. In practice, this batch process approach fits the need of long-term traffic pattern extraction and impact assessment of historical events. For traffic operation engineers, detecting the anomaly in traffic system could alarm them on the events that cause traffic pattern change. For decision-makers, it could help them to quantify the different impacts from historical events and prepare appropriate reaction plan accordingly. For road users, this work could also be extended into an online detection application, which is useful as sending early warnings to road users.

In future work, more corridors could be involved. As running on a long-term historical data, the system anomaly could be easily detected by checking how far it is apart from a normal pattern network. Based on this application, a health monitoring framework for the traffic system can be developed. Future research directions will include: (i) analyze the potential causes of system-level anomaly from real world, then set the priority levels for those real-world events; (ii) summarize the anomalies over a long time and further utilize it to evaluate system-level reliability.

## ACKNOWLEDGEMENT

Our research results are based upon work jointly supported by the National Science Foundation Partnerships for Innovation: Building Innovation Capacity (PFI:BIC) program under Grant No. 1632116, National Science Foundation under Grant No. CNS-1464279 and Iowa DOT Office of Traffic Operations Support Grant. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

## ACRONYM

| | |
|---|---|
| *STPN* | spatiotemporal pattern network |
| *SVM* | support vector machine |
| *UP* | uniform partitioning |
| *MEP* | maximum entropy partitioning |
| *MMP* | maximum migration partitioning |
| *SFNNP* | symbolic false nearest neighbor partitioning |
| *PFSA* | probabilistic finite state automata |
| *AP* | atomic pattern |
| *RP* | relational pattern |
| *MI* | mutual information |
| *SSIM* | structural similarity |
| *HCM* | highway capacity manual |
| *LOS* | level of service |

## REFERENCES

Abdulhai, B., & Ritchie, S. G. (1999). Enhancing the universality and transferability of freeway incident detection using a Bayesian-based neural network. *Transportation Research Part C: Emerging Technologies*, 7(5), 261–280. doi:10.1016/S0968-090X(99)00022-4

Adeli, H., & Karim, A. (2000). Fuzzy-wavelet RBFNN model for freeway incident detection. *Journal of Transportation Engineering*, 126(6), 464–471. doi:10.1061/(ASCE)0733-947X(2000)126:6(464)

Chakraborty, P., Hess, J. R., Sharma, A., & Knickerbocker, S. (2017). Outlier mining based traffic incident detection using big data analytics. Presented at the *Transportation Research Board 96th Annual Meeting*, January 8-12, Washington, D.C. Retrieve from https://trid.trb.org/View/1439336

Jiang, Z., & Sarkar, S. (2015). Understanding wind turbine interactions using spatiotemporal pattern network. *Proceedings of the ASME 2015 Dynamic Systems and Control Conference*, October 28-30, Columbus, OH, USA. doi:10.1115/DSCC2015-9784

Jiang, Z., Liu, C., Akintayo, A., Henze, G., & Sarkar, S. (2017). Energy prediction using spatiotemporal pattern networks. *Applied Energy*, 206, 1022-1039. doi:10.1016/j.apenergy.2017.08.225

Jin, J., & Ran, B. (2009). Automatic freeway incident detection based on fundamental diagrams of traffic flow. *Transportation Research Record: Journal of the Transportation Research Board*, 2099, 65–75. doi:10.3141/2099-08

Jin X., Sarkar, S., Mukherjee, K., & Ray, A. (2009) Suboptimal partitioning of time-series data for anomaly detection, *Proceedings of Conference on Decision and Control*, December 15-18, Shanghai, China. doi:10.1109/CDC.2009.5400158

Kim, J., & Wang, G. (2016). Diagnosis and prediction of traffic congestion on urban road networks using Bayesian networks. *Transportation Research Record: Journal of the Transportation Research Board*, 2595, 108–118. doi:10.3141/2595-12

Li, L., He, S., Zhang, J., & Yang, F. (2016). Bagging-SVMs algorithm-based traffic incident detection. *Proceedings of the 16th COTA International Conference of Transportation Professionals*, July 6-9, Shanghai, China. doi:10.1061/9780784479896.132

Liu, C., Ghosal, S., Jiang, Z., & Sarkar, S. (2017). An unsupervised anomaly detection approach using energy-based spatiotemporal graphical modeling. *Cyber-Physical Systems*, 3(1-4), 66-102. doi:10.1080/23335777.2017.1386717

Liu, C., Gong, Y., Laflamme, S., Phares, B., & Sarkar, S. (2017). Bridge damage detection using spatiotemporal patterns extracted from dense sensor network. *Measurement Science and Technology*, 28(1), 014011. doi:10.1088/1361-6501/28/1/014011

Liu, C., Huang, B., Zhao, M., Sarkar, S., Vaidya, U., & Sharma, A. (2016). Data driven exploration of traffic network system dynamics using high resolution probe data. *Proceedings of 2016 IEEE 55th Conference on*

*Decision and Control* (7629–7634), December 12-14, Las Vegas, NV, USA. doi:10.1109/CDC.2016.7799448

Liu, C., Jiang, D., & Yang, W. (2014). Global geometric similarity scheme for feature selection in fault diagnosis. *Expert Systems with Applications,* 41(8), 3585–3595. doi:10.1016/j.eswa.2013.11.037

Margreiter, M. (2016). Automatic incident detection based on bluetooth detection in northern Bavaria. *Transportation Research Procedia*, 15, 525–536. doi:10.1016/j.trpro.2016.06.044

Rao, C., Ray, A., Sarkar, S., & Yasar, M. (2009). Review and comparative evaluation of symbolic dynamic filtering for detection of anomaly patterns. Sig*nal, Image and Video Processing*, 3(2), 101–114. doi:10.1007/s11760-008-0061-8

Ritchie, S. G., & Cheu, R. L. (1993). Simulation of freeway incident detection using artificial neural networks. *Transportation Research Part C: Emerging Technologies*, 1(3), 203–217. doi:10.1016/S0968-090X(13)80001-0

Sarkar, S., Sarkar, S., Virani, N., Ray, A., & Yasar, M. (2014). Sensor fusion for fault detection and classification in distributed physical processes. *Frontiers in Robotics and AI*, 1, 16. doi:10.3389/frobt.2014.00016

Sarkar, S., & Srivastav, A. (2016). A composite discretization scheme for symbolic identification of complex systems. *Signal Processing*, 125, 156–170. doi:10.1016/j.sigpro.2016.01.018

Sarkar, S., Srivastav, A., & Shashanka, M. (2013). Maximally bijective discretization for data-driven modeling of complex systems. *Proceedings of Americal Control Conference* (2674–2679). June 17-19, Washington D.C., USA. doi:10.1109/ACC.2013.6580238

Solo, V. (2008). On causality and mutual information. *Proceedings of 2008 47th IEEE Conference on Decision and Control* (4939–4944). December 9-11, Cancun, Mexico. doi:10.1109/CDC.2008.4738640

Tang, S., & Gao, H. (2005). Traffic-incident detection-algorithm based on nonparametric regression. *IEEE Transactions on Intelligent Transportation Systems*, 6(1), 38–42. doi:10.1109/TITS.2004.843112

Transportation Research Board (TRB). (2000). *HCM: highway capacity manual.* Washington, D.C., USA: Transportation Research Board.

Wang, Z., Bovik, A., Sheikh, H., & Simoncelli, P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4), 600-612. doi: 10.1109/TIP.2003.819861

Weather Underground (2017, October 10). Weather history for KDSM. Retrieved from: https://www.wunderground.com/history/airport/KDSM/

Wibral, M., Rahm, B., Rieder, M., Lindner, M., Vicente, R., & 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. doi:10.1016/j.pbiomolbio.2010.11.006

Wu, L., Liu, C., Huang, T., Sharma, A., & Sarkar, S. (2017). Traffic sensor health monitoring using spatiotemporal graphical modeling. *Proceedings of the 2nd ACM SIGKDD Workshop on Machine Learning for Prognostics and Health Management*, August 13-17, Halifax, Nova Scotia, Canada.

Yao, B., Hu, P., Zhang, M., & Jin, M. (2014). A support vector machine with the tabu search algorithm for freeway incident detection. *International Journal of Applied Mathematics and Computer Science*, 24(2), 397–404. doi: 10.2478/amcs-2014-0030

Yuan, F., & Cheu, R. L. (2003). Incident detection using support vector machines. *Transportation Research Part C: Emerging Technologies*, 11(3), 309–328. doi: 10.1016/S0968-090X(03)00020-2

## BIOGRAPHIES

**Tingting Huang** is a Ph.D. Candidate in Transportation Engineering at Iowa State University. She is working at Institute for Transportation as a graduate research assistant since 2014. She mainly focuses on big data analytics in transportation operation and safety area. Huang's research interests include data quality control, traffic state estimation, signalized intersection analysis, highway safety and work zone operation analysis. She is served as a reviewer for several committees in Transportation Research Board.

**Chao Liu** is a research assistant professor at Department of Energy and Power Engineering, Tsinghua University, Beijing, China. Previously, he was with the Iowa State University as a Postdoctoral Fellow. He received the B.Sc. degree from Huazhong University of Science and Technology, Wuhan, China, in 2008, and the Ph.D. degree from Tsinghua University, Beijing, China, in 2013. Dr. Liu's research interests include structure dynamics, machine learning, diagnosis, prognosis, and health monitoring.

**Anuj Sharma** is an associate professor in the Civil Construction and Environmental Engineering Department at Iowa State University. He also holds a joint appointment as a Research Scientist with the Institute of Transportation. In these positions, he teaches transportation engineering courses to undergraduate and graduate civil engineering students, conducts research in the transportation operations area, and participates in numerous professional organizations. Dr. Sharma is currently leading the

REACTOR (REaltime AnalytiCs of TranspORtation data) laboratory. The lab can ingest multiple streams of real-time data to assist in driving transportation policy decisions. The efforts are focused on ingestion, real-time analytics, batch processing, visualization/front end development, and archiving of numerous data streams.
He coauthored more than 94 peer-reviewed publications including 31 journal papers, 1 book chapters and one patent. He has also served as a reviewer and session chair for several technical journals and conferences.

**Soumik Sarkar** is an assistant professor of Mechanical Engineering at Iowa State University. Previously, he was with the United Technologies Research Center for 3 years as a Senior Scientist. Dr. Sarkar's research interests include Statistical Signal Processing, Machine Learning, Sensor Fusion, Fault Diagnostics and Prognostics, Distributed Control and Complexity Analysis with applications to complex Cyber-Physical Systems such as aerospace, energy and smart building systems, transportation, manufacturing and agriculture systems. He coauthored more than 100 peer-reviewed publications including 36 journal papers, 4 book chapters and one magazine article. He has also served as a reviewer and session chair for several technical journals and conferences. Dr. Sarkar is currently serving as an Associate Editor of Frontiers in Robotics and AI: Sensor Fusion and Machine Perception journal.