Traffic Congestion Prediction System using Artificial Pattern-based Dynamic Time Warping

Paper Identification Number: SCS12-012
Naphat THEERAWAT[1], Kittipong HI-RI-O-TAPPA[2A], Suttipong THAJCHAYAPONG[2B]

[1]Faculty of Information and Communication Technology
Mahidol University
NakhonPathom, Thailand
E-mail: naphattheerawat@gmail.com

[2]NECTEC, National Science and Technology Development Agency
Bangkok, Thailand
[2A] E-mail: Kittipong.h@gmail.com
[2B] E-mail: Suttipong.Thajchayapong@nectec.or.th

Abstract

This paper proposes a real-time congestion prediction system for highway segments. The system is derived from an incident detection algorithm proposed in [1]. The key advantage of the proposed system is its adaptability which is achievable through the use of Dynamic Time Warping with the artificial patterns derived from the optimization of the hierarchical clustering of the congested highway patterns. With these artificial patterns, the proposed system requires neither calibration nor site-specific training dataset and can immediately be deployed to assess traffic congestion at a new site that has similar traffic characteristics. The ability of the proposed system to generically model traffic patterns prior to congestion presents the main advantage over many previously proposed approaches which need to be calibrated with training dataset when deployed at a new site. Furthermore, the conducted performance evaluation using real-world data shows that the system can achieves over 82 percent detection accuracy while the false alarm rate remains low.

Keywords: Congestion Prediction System, Dynamic Time Warping

1. General Introduction

In Thailand, there are many algorithms sharing the objective of making highway traffic congestion prediction in real-time, yet few of them have accomplished the task without using site-specific training dataset. Calibration is a labor-intensive job requiring massive amount of effort and determination to modify the algorithm until the performance reaches the acceptable level; having to do such task repeatedly on highway data could be time consuming and unbearable since there are a growing number of highways in every country.Therefore, coming up with a paradigm that does not require site-specific dataset and can be immediately deployed at a new site would definitely be a very pleasant solution formany researchers. Building up from the idea, the system being introduced solely uses artificial patterns achieved from the optimization of the hierarchical clustering of the congested highway patterns as proposed in [1] along with Dynamic Time Warping to forecast the highway traffic status.

2. Background

2.1 Dynamic Time Warping

Dynamic Time Warping (DTW) is an algorithm that is famous for its capability of determining the similarity between two patterns. It was first proposed by Berndt et al [2] in 1994 and has been applied into several fields of studies including Computer Science such as image processing, speech recognition, etc. as a warping axis is introduced to helps it perform better than Euclidean distance. However, it has never been introduced into the field of transportation engineering. In this paper, this algorithm pays an important role as it is used to find the similarity between time series of microscopic traffic variables and the artificial patterns.

Figure 1 Euclidean Distance

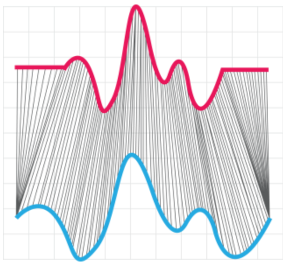

Figure 2 Dynamic Time Warping

2.2 Traffic Flow Theory

Traffic congestion is the state in which the flow of vehicles becomes slower than usual due to the occurrence of event that is either predictable (recurrent) or unpredictable (non-recurrent). It is found that if the congested state can be foretold for a single time unit, the length of the congestion will be reduced by 1.5 to 3 times [5].

To forecast traffic state, the proposed system uses three traffic indicators as proposed in [1, 3] as the manipulation of these values expresses the change in traffic state. The three indicators are listed as follow:

2.2.1 Speed

Traffic status can be derived from speed as the value that falls under the threshold represents the presence of traffic congestion.However, speed alone is highly sensitive data and can be affected by many other factors such as the amount of vehicles, road characteristics, driver's preference, thus it should be considered along with other variables to indicate the state of traffic.

2.2.2 Standard Deviation of Speed

Standard deviation of speed indicates the difference between speeds in the period of interest. The high value of this variable signifies that there has been a lane blockage caused by

either accidents or traffic incidents which often lead to congestion.

2.2.3 Time headway

Time headway is the time gap between vehicles. Its value along with speed can be used to determine whether an accident would occur; for instance, if the time headway is lower than the human response rate while the speed is high, the likelihood of the occurrence of an accident is presumably high.

3. Methodology

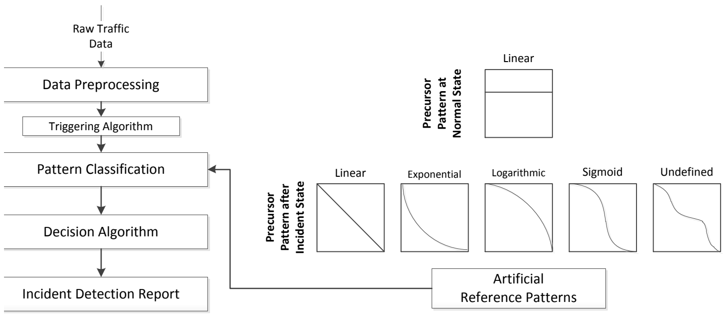

Figure 3 shows the System Framework [1]

3.1 Data Preprocessing

The nature of raw traffic data is noisy and inappropriate to be used right after its presence in the repository. Generally, the input dataset consists of noises and, sometimes, discontinuity, therefore it should be fixed before processing, especially when the system requires a continuous dataset. The preprocessing steps are as follow:

3.1.1 Data Interpolation

To maintain the continuity of microscopic traffic variable time series, the absence of data must be eradicated by simulating and filling data into the gap.

3.1.2 Data Smoothing

After the interpolation, the continuity has been restored; however the dataset still cannot be used as it contains noises which can reduce the prediction accuracy. Noise removal is achievable by averagingout data for each interval.

3.2 Data Processing

3.2.1 Pattern Classification& Artificial Reference Patterns

The system uses artificial patterns derived from [1]; the patterns can be categorized into two main types which are congested pattern and uncongested pattern as shown in figure 4.The congested pattern is further subdivided as the decrease of speed tends to occur in different forms; for instance, linear reduction, exponential fall, etc. The pattern classification is done through the use of Dynamic Time Warping to find the similarity between the input and the artificial patterns.

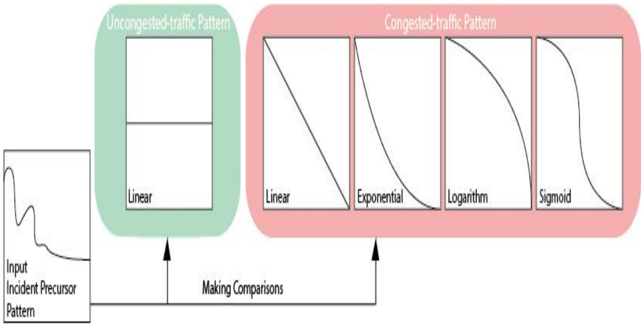

Figure 4 illustrates the comparison between inputted pattern and Artificial Traffic Congestion Patterns

3.2.2 Decision Algorithm

The outcome of the forecast will be yielded based on the comparison. If the inputted dataset bears resemblance to any of the congested patterns, the system will foretell the traffic status as congested.On the contrary, if the dataset shares more similarity with the uncongested pattern, the result of the forecast will be opposite.

3.2.3 Incident Detection Report

For analysis and future usage, the result of the prediction is stored into a log file and in our system database.

4. Performance Evaluation

The performance evaluation was conducted using CCTV images as benchmark to assess one-week data measured at Daokanong-Suksawat expressway from June 14th, 2012 to June 21st, 2012. To determine the quantity of data required for correctly identifying the accuracy of prediction, we categorize the results; each of them uses different amount of data during the processing step. To evaluate the effectiveness of the system, Detection Rate (DR) and False Alarm Rate (FAR) are used as the measures to assess the accuracy [4] of the traffic congestion forecast.

As seen from the result of the performance assessment shown in table 1, the detection rate rises as the amount of data used increases, whereas the False Alarm Rate falls correspondingly.

Table 1 shows the performance of the proposed system when different quantity of data is used

Parameter	Amount of data used			Result of the performance stated in [1]
	5-minute data	10-minute data	15-minute data	
Detection Rate	69.67 %	75.48 %	82.53 %	91 %
False Alarm Rate	30.32 %	24.51 %	17.46 %	25 %

5. Conclusion

This paper introduces a system implementing the concept proposed in [1]. The system preprocesses microscopic traffic variable time series to get rid of discontinuity and noises by performing interpolation and smoothing. Afterwards, comparison between the inputted dataset and artificial patterns are made through the use of Dynamic Time Warping to determine the similarity between patterns; the most resemble pattern will be considered as the appropriate outcome which will then be yielded and stored into a log file and the database for evaluation and analysis.

- Due to the absence of sunlight, the raw traffic data produced by the image processing process of the highway camera sensors becomes inevitably unreliable during night hours; this introduces discontinuity to the dataset which needs to be preprocessed before used.

- Once deployed, thetradeoff arises as the accuracy of traffic congestion prediction is taken into consideration along with processing time. The precision of the forecastdepends on the amount of data used, but the higher quantity of data will result in the slower processing time.
- As a result of using artificial patterns along with Dynamic Time Warping to make prediction, the system needs neither site-specific training dataset nor calibration, and the resources such as budget and time used for calibration can be cut off.

6. Reference

[1] Hi-ri-o-tappa, K. &Thajchayapong, S., "Generalizability and Transferability of Incident Detection Algorithm Using Dynamic Time Warping," ITS World Congress, Vienna, Austria
[2] Berndt, D. & Clifford, J. (1994), "Using dynamic time warping to find patterns in time series," AAAI-94 Workshop on Knowledge Discovery in Databases (KDD-94), Seattle, Washington.
[3] Hi-ri-o-tappa, K., et.al, "A Novel Approach of Dynamic Time Warping for Short-Term Traffic Congestion Prediction," Transportation Research Board of the National Academies, Washington, D.C., 2011.
[4] Parkany, E. and C. Xie, "A Complete Review of Incident Detection Algorithms and Their Deployment: What Works and What Doesn't," The New England Transportation Consortium, 2005.
[5] Zhang, L., and D. Levinson, "Some Properties of Flows at Freeway Bottlenecks," Transportation Research Board: 83th Annual Meeting 2004.