A Large Incident Management System (IMS)
For Intelligent City Development

Project Plan
Version 1.0
September 30, 2018

Supervisor:
Dr. Reynold C.K. Cheng

Group members:
LI Jianxiang (3035235214)
LEI Haoyu (3035235044)
Table of Contents

Table of Contents ................................................................................................................................. 1
Revision History ................................................................................................................................. 1
1. Background .................................................................................................................................. 2
2. Objective ...................................................................................................................................... 2
3. Methodology .....................................................................................................................................
   3.1 Methods to be Implemented .................................................................................................. 3
   3.2 Data Preprocessing ................................................................................................................. 4
   3.3 Experiment Metrics .................................................................................................................. 4
   3.4 Visualization .............................................................................................................................. 5
4. Schedules and Milestones ............................................................................................................. 6
References ............................................................................................................................................... 7

Revision History

<table>
<thead>
<tr>
<th>Date</th>
<th>Recent for Changes</th>
<th>Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018.9.30</td>
<td>Initial document</td>
<td>1.0</td>
</tr>
</tbody>
</table>
1. Background

In recent years, Hong Kong government has planned to make the city more intelligent, and traffic is one of its main focus. As the gateway to mainland China and a crucial financial hub in Asia-Pacific region, Hong Kong bear enormous amount of traffic on a daily basis. According to Hong Kong Transport Department, about 780,000 vehicles are currently running in the city, which is a significant figure compared with its road length of 2,101 kilometers [1]. Such a large vehicle base and limited road capacity requires more efficient use and allocation of transport resources. Additionally, the large traffic flow happening in Hong Kong produces a significant number of incidents and retards the traffic by a large extent.

At present, incidents on the road is monitored by Traffic Control Center, and the process is manual. A large incident management system, with algorithms automatically detecting traffic incidents, will free more manpower and make the incident monitoring process more efficacious. An accurate incident detection system can not only benefit Transport Department on traffic management such as traffic flow control by making use of police power and intelligent traffic lights, but also helps individual traffic participants to have a more accurate estimation on the travel time.

2. Objective

In this project, we are aiming to find an optimal algorithm allowing accurate detection of traffic incidents. To meet such a goal, we will explore, devise, implement and test algorithms that are suitable to the nature of Hong Kong traffic data. We will then select the optimal algorithm according to several metrics: mean time to detect, false alarm rate, and detection rate. A detailed explanation of these metrics will be provided in section 3.

The intermediate objective for the project is to develop a demo showcasing how an incident detection algorithm works, preferably with visualization aid such as a web page. In this demo, we will demonstrate a use case where an incident happens and gets detected.

The final goal for the project is to provide a report offering an optimal algorithm for incident detection, and an exhaustive analysis discussing how it fits the nature of Hong Kong traffic condition.
3. Methodology

Multiple methods are proposed to solve the incident detection problem. As different models usually obtain different performances on different datasets with various nature, a series of experiments would be done on some existing incident detection methods. These methods include pattern matching [2, 3], deviation detection [4, 5, 6] and supervised machine learning models [7, 8, 9, 10, 11]. Through implementing and testing the given methods, the strength and weaknesses of each method would be analyzed, and the results would be displayed in an easily understandable visualized interface.

3.1 Methods to be Implemented

There are three categories of method that is in our interest, namely pattern matching, deviation detection, and supervised models.

- **Pattern Matching: California [2], DELOS [3]**
  This method compares the difference of traffic flow features between upstream and downstream of the incident location. In other words, the model estimates the disparity of traffic before a certain location and after a certain location to see if the difference exceeds a certain threshold, and therefore detected as an incident. For example, if there is a significant speed drop after vehicles passed a certain road section, the model detects it as an incident.

- **Deviation Detection: Time Series [4, 5], Topic Model [6]**
  This method compares the difference of current speed distributions of the traffic flow with the usual speed distribution of the estimated time and location. If the current speed at a specific road segment is significantly different with the usual speed distribution on the road segment at the same hour of the day, the segment is marked as incident.

- **Supervised Models: SVM [7], NN [8, 9, 10], Deep Learning [11]**
  Support Vector Machine (SVM), Neural Network (NN) and Deep learning are some common models practiced in the field of machine learning. They are called supervised models because we need to train and tune the models using labeled data to make them able to do classification. The detailed implementation and theories will be explained in further deliverables.
3.2 Data Preprocessing

Currently, there are two datasets available for the project. One of them is a confidential dataset of Hong Kong taxi GPS locations, the other is an open dataset of California highway I880.

There are two major potential challenges hidden in Hong Kong dataset, inaccuracy and sparsity. Multiple approaches will be made to compensate for the detection accuracy loss caused by the specific nature of dataset, such as selecting location and time with higher density of data points and fixing the data point locations using map matching algorithms. With the support of visualization methods (discussed in 3.4), we will have a more intuitive understanding of the dataset, which helps with finding the parts of dataset with higher density and accuracy.

California dataset provides clean, accurate and dense data from fixed highway sensors. This dataset has different characteristics with the GPS location of individual vehicles as it provides average stats of the whole traffic flow. It serves the project with two major purposes: testing the code validity and analyzing the strength and weaknesses of different methods on different kinds of dataset.

As the quality of data also largely influences the performance of different models, more data cleaning and preprocessing works will be done together with the experiments. Furthermore, we will leverage other sources of data to aid experiments if necessary.

3.3 Experiment Metrics

After fitting our datasets into these existing models and implementation, we will do experiments on different parameters and different circumstances such as separating night and daytime, urban roads and highways. A set of applied metrics will be used in this project, which is detection rate (DR), false alarm rate (FAR) and mean time to detect (MTTD). DR measures how many incidents are successfully found, whereas FAR measures how many proportions of common traffic conditions are falsely identified as incident. MTTD provides how much time on average is needed to detect an incident.

$$\text{Detection Rate} = \frac{\text{Detected \# of Incidents}}{\text{Total \# of Incidents}}$$

$$\text{False Alarm Rate} = \frac{\# \text{ of observations falsely considered as incident}}{\# \text{ of observations that have no incident}}$$

$$\text{Mean Time to Detect} = \frac{\sum \left( T_{\text{detection}} - T_{\text{incident}} \right)}{\# \text{ of detected Incidents}}$$
3.4 Visualization

In order to help human better understand the performance of the models, several data visualization methods will be applied. For instance, as detection rate and false alarm rate are often in a trade-off relationship, the curve plotted on detection rate and false alarm rates will be generated once selected a model, parameters and certain categories of the data to apply.

![Figure 1. A visualization of the trade-off between detection rate and false alarm rate in topic model](image1)

Besides charts and diagrams, an incident map demonstration will also be presented. The visualized map will display the vehicle data points on the map and the estimated incident location and time by our methods, comparing with the real incident.

![Figure 2. A visualization of taxi GPS coordinates sampled from Hong Kong dataset, labelled by red points](image2)
4. Schedule and Milestones

<table>
<thead>
<tr>
<th>Time</th>
<th>Work</th>
<th>Deliverables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present – Nov.30th 2018</td>
<td>Implement existing models and experiment on our own dataset</td>
<td>Baseline report on experiment results and implementation code</td>
</tr>
<tr>
<td>Dec.1st - Dec.31st 2018</td>
<td>Prepare a demo for implemented models</td>
<td>Demo code or software</td>
</tr>
<tr>
<td>Jan.1st - Jan.31st 2019</td>
<td>Continue research and propose methods to improve algorithms, and prepare interim report</td>
<td>Interim report</td>
</tr>
<tr>
<td>Feb.1st – Mar.31st 2019</td>
<td>Improve existing method and develop our own model; draft final report</td>
<td>Detection code or software</td>
</tr>
<tr>
<td>Apr.1st – Deadline of FYP</td>
<td>Finish final report</td>
<td>Final Report and Implementation of refined model</td>
</tr>
</tbody>
</table>
References


