2014

Estimate status of roads:Using GPS Bus tracking data

Tran Tin
Ton Duc Thang University, tttin@it.tdt.edu.vn

27R. Estimate status of roads:Using GPS Bus tracking data.

Tin, Tran
Ton Duc Thang University
tttin@it.tdt.edu.vn

Abstract

Assess the traffic situation will be very good for finding the fastest route or avoid the traffic jam. In order to do so, we need to know the traffic situation on all the roads. Speed of the flow of vehicles on the road can be measured by number of vehicles go by a check point within a unit of time. This measurement shows the status of that road, a low value points out a jam and a high value let us know the road is clear. But in the case there are no private lanes for every kind of vehicles and vehicles are moving straggling, counting vehicles is difficult. Another method is monitoring the flows though camera traffic and make evaluation of the status. This also requires more staff and the evaluation is not continuously nor homogeneous.

Nowadays, all buses are equipped with GPS real-time trackers which send their data of speed, location to the center control. In the case buses is moving together with other vehicles on a busy road, we assume that buses having an average speed that could represent the movements of other vehicles. Within a specific road, there're many buses go by within a period of time, making a specific set of speed data. From these sets, we can estimate average speed of this road. We also calculate the reliability of that result with some statistical functions. Furthermore, a huge amount of data sending to the center control is difficult for data processing. The system need to be designed in parallelism architecture for performance improving. The data will be divided by a receiver before they are sent to their sub-system for the estimation. After all, data collected day by day could be used to verify the current day's calculating. So that, an effective database design is the last but not least.

Moreover, the data collected day by day will provide the behavior of the flow on a road. Which region usually have jam or when the roads will be more heavier can be known by review the estimate result in the past. So that, the GPS data from buses will be more useful if they are collected and processed together. This paper shows how to collect speed data from buses, make sets of data and from that estimate the status of roads. By comparison with other methods, we also verify the GPS data from buses is useful or not.

Keywords

GPS tracker, Traffic control, Speed of traffic flow.

1. Introduction

Monitoring the flow of traffic is a very important subject for a city management. There're many kind of vehicles that are trucks, buses, cars and bikes moving rapidly on the network of roads. With some handheld's applications using built-in GPS; the drivers can send their speed and location data to the center control. We cannot know the speed of all vehicles, but with acceptance rate of number of data received on number of vehicles, the center control can estimate the status of roads.

Everyday, there're many buses go around the city. Tracker is mounted inside every bus, and they will send to the center control data of its speed and location every few second. Within a period of time and a specific road, we could get a set of data of speed. From that set of data, we can calculate the average speed that represents the status of the observed road. The reliability of our calculating depends on how many buses have sent out their data, how long an observation took and the length of those roads.

We also represent an architecture that will be needed to make our system real-time. When monitoring all roads in the city and all running buses, a huge amount of data could be sent to the center control and that data need to be processed as fast as possible. So that, the database of roads should be split into junior databases by an adaptive method and the data receiver need to distribute data to the right region which the data belong to. Furthermore, this architecture can make the system parallelism when we implement individual calculator for every junior database.

On the other hand, with recorded average speed, we can observe how the speed of a specific road changes by time, or learn the relation between adjacent roads by viewing their speed values at the same time. The last but not least, in the case some roads have received no data, the system should review the database and make estimation from information in the past or from adjacent roads' status.

2. Related works.

2.1 The buses tracker system.

GPS (Global Positioning System) is a system that tells the users their current location on the earth's surface. The GPS device also gives out the value of current time. A location is represented by a geographic point which has a pair of value: latitude and longitude. When the users move some distance, their location change corresponding. And from the difference of moving time and the distance, users can know their speed that is instantaneous velocity of this moment. As [1], the value of speed from GPS device is acceptable as the one from odometer.

To monitor a bus, a tracker is mounted. While moving with the bus, the tracker record time, speed, location and some other contexts. The tracker will send that information to the center control though wireless networks. At the center control, operators know where the buses are by mapping received locations into a database of digital map.

2.2 The mathematic function.

As the introduction mentioned above, we could receive many speed data from many buses after a time-slide in a specific segment. Of course, they are not the same value. According to [2], mean speed of that buses flow can be calculate for spatial and temporal regions.

$$\overline{\mathrm{v}}_{\mathrm{s}}(spatial) = \frac{1}{n}\sum_{i=1}^{\mathrm{n}} \mathrm{v}_{\mathrm{i}}$$

$$\bar{v}_t(temporal) = \frac{n}{\sum_{i=1}^{n} \frac{1}{v_i}}$$

The statistical sample variance defined as follow:

$$\sigma_s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (v_i - \bar{v}_s)^2$$

3. The architecture of the system

The system has three main components: a database with all of roads' segments that we need to monitor the speed. A data receiver processes all the information from GPS trackers, and determines which segment a piece of information belongs to. After a period of time and all the information we've observed, the data analyst component should give out statements about status of roads' segments.

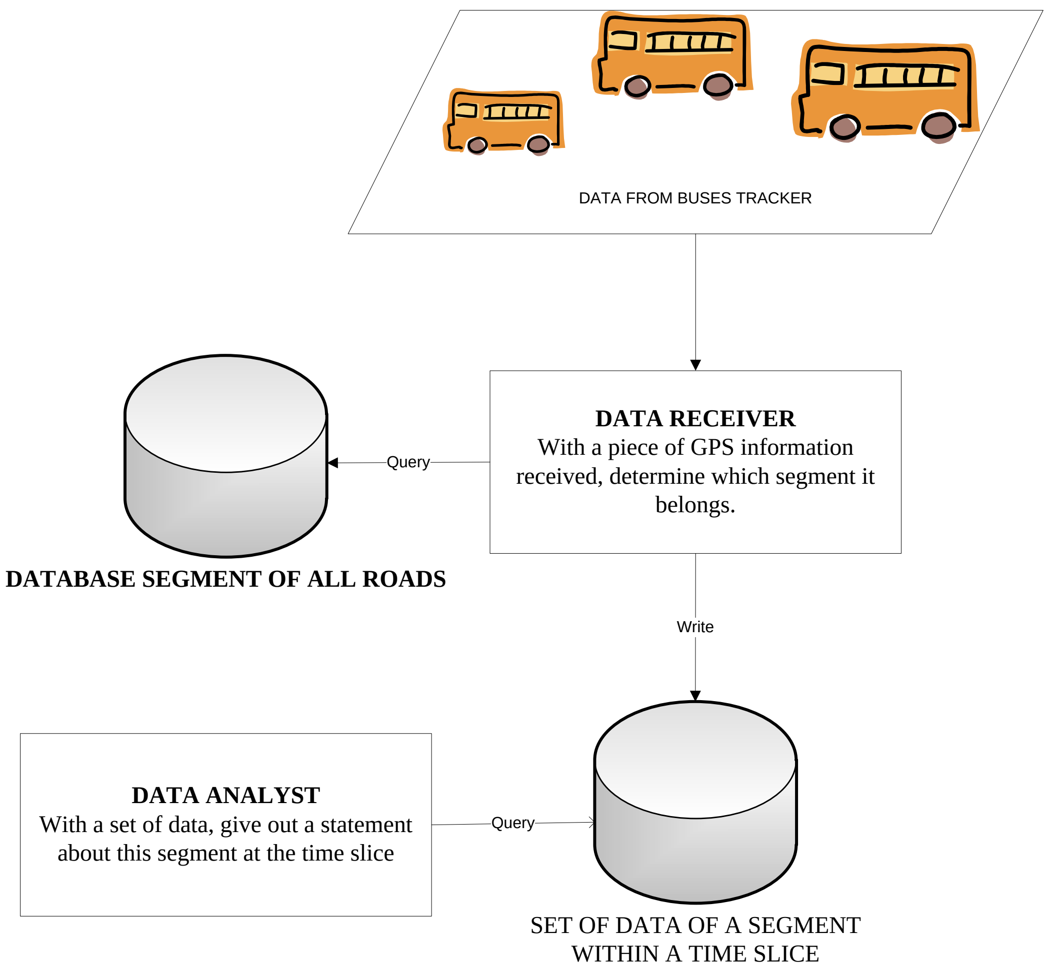

Figure 1: An overview on system architecture.

3.1 Database of segments of all roads.

The database is a model that represents all the roads in the city, and stores the calculated speed of these roads time by time.

In fact, roads always have curves and unstable features, for example, the width. So that, a road may be split into many segments. The size of a segment is variable but need to be considered. If a segment is too long, the observation will not be reliable. In contract, a very short segment will get few data; even no data will be recorded. With the dimension of time, the traffic is changed rapidly by time; we need to make a suitable time-slice.

A segment is represented by a rectangle having a specific width, and stored by geographical location. A speed data is always enclosed with location data (also a geographical point), and can be verify which segment it belongs to.

After a time-slice of observation, a segment can receive some speed data from some buses (and also can not) and make a set of information. And the calculator will work with that set to get an average speed. A small value of average speed shows that most of buses within this segment are moving slowly. And if these buses move fast, we would get a bigger value.

When make a new observation of the next time-slice, the old result should be stored in the database for data-mining purposes.

3.2 Data receiver component.

Two main objects in this step is skipping invaluable information and make a set of data for each segment. Not all of information is useful. There is some bad information from trackers, for example, a incongruous speed, or speed change so fast, or the information is out of observed segments.

A piece of information need to be known which segment it belongs to, in real world it means where the piece of information has been recorded. After a period of time, this component gives out a set of recorded speed for each segment. There're some segments having no record.

Otherwise, if the main database of street is divided into junior database. For example, the center control wants to separately monitor downtown area and others. This component has responsibility of sending a set of data to the right region.

3.3 Data analyst component.

There're two cases with a segment after a time-slice of observation.

In the case we've gotten a set of data, we apply a function to estimate the speed. For example, we could use average function, median value or harmonic function. The reliability depends on standard deviation, the force of this data set and the stabilization of the road.

On the other hand, if the set of data is empty, the system will review the history of the same segment, or some adjacent segments that belong to the same road for estimation.

4. Implementation of the system

4.1 Database of segment.

A segment is represented by a start geographical node and an end one. A segment has fixed width. The width is different from other segments and can be configured in the database. We can also calculate the value of orientation of a segment. With a speed data inside a segment but different orientation's value, that data could be ignored.

For continuously monitoring, the segments should be consecutive division. It means the end point of a segment is a start point of another one. At junction or roundabout, traffic flow is very sophisticated and need a special processing.

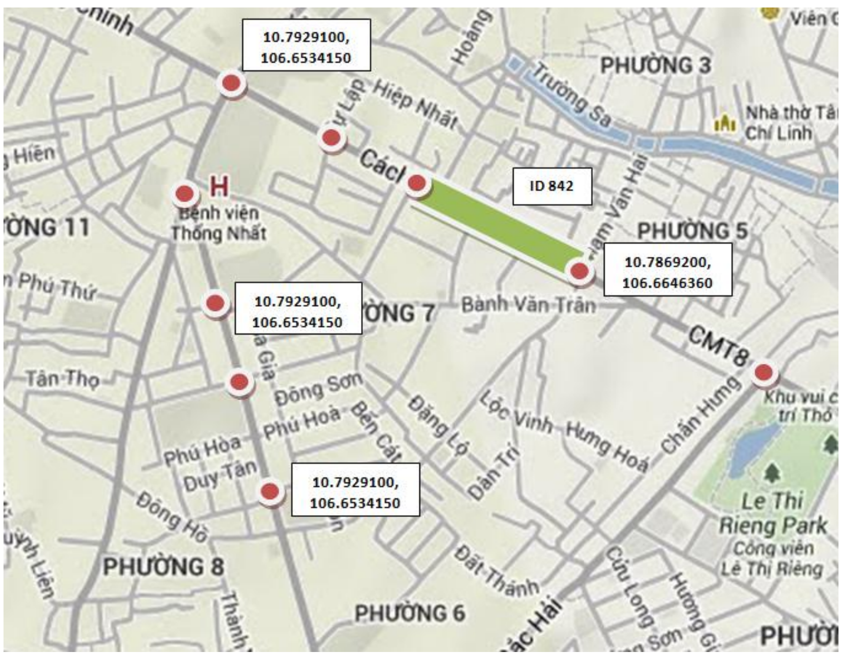

Figure 2: An example of segment division

4.2 Division of observation region.

When a bus sends a piece of speed data, the receiver needs to know which segments that contain that speed data. The database has thousands of segments, so that queries are numerous. On the other hand, when the city need more one control center and one center take control a region; we need to divide the database.

With the database, division can be applied by latitude and longitude. The map of city is now divided into many equally rectangles. So that, segments that belong to the same rectangle area will be stored in the same table in the database.

When the system works, with a speed data received, the system can identify which rectangle area that data belongs to and query only the correlative table to find the segment which contains the data. It helps reduce the number of queries in the database. Instead of checking all the segments in the database, we need to check only one table that is smaller than the whole database many times.

5. Conclusions

This paper has provided a brief view of a system which can estimate the status of roads using the GPS trackers' data. This method is cheaper than others, for example, staff to counting vehicles or traffic cameras system. When the city has more buses, and the database recorded a large historical moving speed of all buses, the effectiveness of the system will increase.

However, when the buses is moving in a private lane, or new roads are built, the system need to reconfigure and all the history of related segments are unusable. Beside introduced mathematic formulas, the system could be improved by some custom function that can be point out for the history of estimated speeds.

References

[1] Wenjuan Zhao, Anne Goodchild, Edward McCormack (2010) "Estimating truck travel speed from GPS spot speed" University of Washington.

[2] Sven Maerivoet, Bart De Moor (2008) "Traffic Flow Theory" Department of Electrical Engineering ESAT-SCD (SISTA) of Katholieke Universiteit Leuven