International Journal of Internet, Broadcasting and Communication Vol.7 No.1 6-9 (2015)
http://dx.doi.org/10.7236/IJIBC.2015.7.1.6

IJIBC 15-1-2

Improving Utilization of GPS Data for Urban Traffic Applications

Nguyen Duc Hai[1], Nguyen Tan Phuc[1], Doan Khue[1], Ta Ho Thai Hai[1], Pham Tran Vu[1], Huynh Nam[1] and Le Thanh Van[1]

[1]*Faculty of Computer Science and Engineering*
Ho Chi Minh City University of Technology, Vietnam
ltvan@cse.hcmut.edu

Abstract

The use of Intelligent Transportation System (ITS) is promising to bring better solutions for managing and handling the city traffic. This system combines many fields in advanced technology such as Global Positioning System (GPS), Geographic Information System (GIS) and so on. The basement of applications in ITS is the effective collections and data integration tools. The purpose of our research is to propose solutions which involve the use of GPS time series data collected from GPS devices in order to improve the quality of output traffic data. In this study, GPS data is collected from devices attached to vehicles travelling on routes in Ho Chi Minh City (HCMC). Then, GPS data is stored in database system to serve in many transportation applications. The proposed method combines the data usage level and data coverage to improve the quality of traffic data.

Keywords: *Intelligent Transportation System, time series data, Global Positioning System.*

1. Introduction

GPS data from city buses is a valuable resource from urban traffic application. It can be used for many different purposes, such as monitoring the current status of the bus systems, helping commuters to catch a bus at the right location and time, monitoring city traffic, detecting congestions. However, as the quality of GPS devices are varied, the GPS data may be too noisy. Consequently, the actual utilization of GPS data collected is low. The GPS data is collected from Ho Chi Minh City (HCMC) bus system is suffering from this problem. Even though large amount of GPS data can be collected in real-time, in our application, which shows the current state of HCMC traffic status in real-time, only a small fraction of the data is usable.

In this paper, we first explore the quality of the bus data collected from HCMC bus system. Then, a method of correcting GPS signal based on bus routes is proposed. The experimental result shows that the utilization of GPS data can be improved significantly after correction.

2. Evaluating GPS data

In our current system, GPS data is collected from GPS devices attached on buses, the signal will be sent to the Bus Operation Centre of HCMC. Then the data will be transmitted to our processing server in real-time. The server also receives GPS signal from mobile devices such as smartphones. A simple statistical calculation shows that there are about 3.5 million GPS signals which the system receives per day in which most of the data is from buses, as in Figure 1. However, the ratio of useful data is only about 17%.

Manuscript Received: Nov. 15, 2014 / Revised: Jan. 5, 2015 / Accepted: Feb. 2. 2015
Corresponding Author: ltvan@cse.hcmut.edu
Tel: +84-8-38-647-256, Fax: +84-8-38-653-823
Faculty of Computer Science and Engineering, Ho Chi Minh City University of Technology, Vietnam

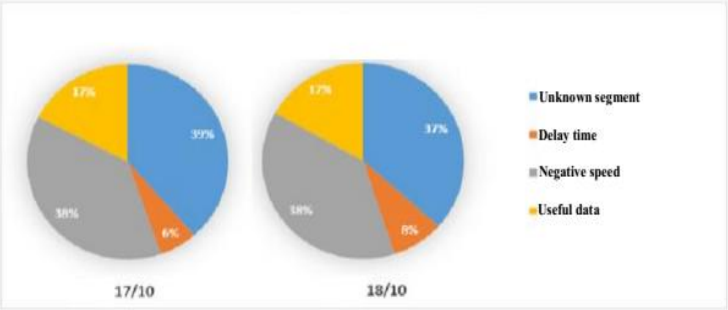

Figure 1. The ratio of data usage

Figure 1 shows the classification of errors in data collection, the orange piece expresses 17% data in use. The errors are normally caused by the device errors, the lost of data package in transmission (especially, with UDP protocol), the delay time, the recorded time error, the change of traffic infrastructure affecting the matching mistake to city map. Moreover, the data received in [9:00AM to 10:00PM] is smoother than [4:00AM to 5:00PM]. This also affects the coverage of data in analysis. In general, the errors can be classified in two groups: the spatial error and the temporal error.

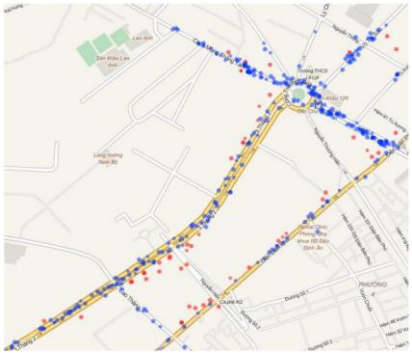

Figure 2. Simulation the usage and coverage of data

Figure 2 shows the state of GPS data from 10:00AM to 11:00PM in specific area in HCMC, the green points express the useful data and the red points express the removable data as it is nonsense in determining the route segments.

3. Improving utilization of GPS data

3. 1 The Improvement of Data Usage

The study mentions methods to solve three issues of GPS data in current system: *the speed with negative value (38%), the delay time (6% - 8%), the confusion in determining specific segment (37% - 39%).*

- Zero speed: The most of GPS data which has the velocity with zero value recorded in the interval time 22:00 to 04:00 or at the main station, the removal of them is valid to avoid much noisy data in applications.
- Delay time: To prevent the deviation in time between receiver and sender, the research uses NTP server to synchronize the timestamp for all of them, this way also reduces the time error in recording the GPS data.
- Unknown segment: This issue is classified to spatial error. The segments are described by rectangle form, the GPS point is considered to belong the segment if it is in domain of segment or close to the boundary of rectangle in enable threshold (available threshold is 5 meters). The Figure 3 expresses the way to determine whether the GPS point belongs the segment.

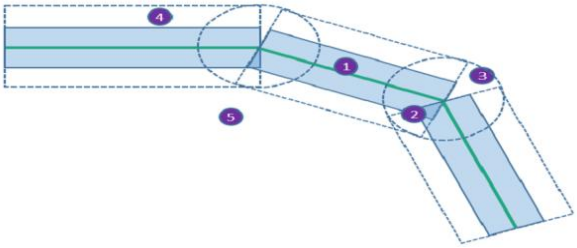

Figure 3. The determination of unknown segment

By calculating the distance from GPS point to vertices of extend segments, the proposed method can check more GPS points (point 3) which belong segment to increase the amount of useful data.

3.2 The Improvement of Data Coverage

The approach is based on the route of vehicles attached GPS device to fulfill the data by creating new virtual points on segment went through. Using the map of route [1] can help not only to predict the next segment in motion but also determine the segment where mediate GPS points went in time series data.

Figure 4. The forecast of mediate segment based on route map

The Figure 4 illustrates a specific circumstance, the green points are GPS signal received by a bus, and these points belongs route of bus expressed by red line, hence it completely sure that the motion of bus will run on reasonable route.

4. Experimentation

For evaluating and validating the proposed methods, some experiments were carried out on GPS data collected in October, 2014. The Table 1 shows the result in finding segments for GPS data point in seven days, from 18-10-2014 to 24-10-2014.

Table 1. Result of determining segment

Date		18-10	19-10	20-10	21-10	22-10	23-10	24-10
Amount of data record		3169646	3561852	3559515	3480681	3483431	3476367	3502442
Old method	Segment	846676	939422	1050127	1007350	1020485	1002670	1013023
	Percent (%)	26.71	26.37	29.5	28.94	29.3	28.84	28.92
New method	Segment	839999	932253	1042326	999440	1013462	995475	1006066
	Percent (%)	26.5	26.17	29.28	28.71	29.09	28.64	28.72
Ratio of improvement		0.21	0.2	0.22	0.23	0.21	0.2	0.2

The new method has high evaluation in effectively decreasing amount of unknown segments, which means the data contributed in our ITS applications will be increased about 0.2%.

The Figure 5 shows the ability of enhancement in fulfilling the signal on segments that have not received GPS data. The orange line expresses the new method and the blue line is considered the current implementation (used in Smart Traffic application located at http://traffic.hcmut.edu.vn). After processing 700 raw GPS data, the current method returns the result with 63 segments (about 9%) whereas the proposed method outputs 108 segments. Therefore the new algorithm brings out the high effectiveness in increasing the coverage level of data on segments.

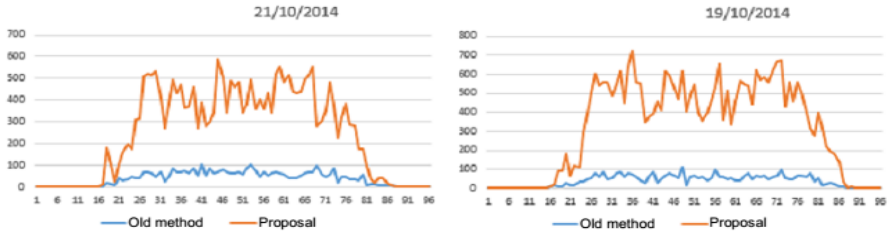

Figure 5. The coverage of data

5. Conclusion

This research has brings out the current view of GPS data collected from GPS devices in HCMC. Following result in the study, two methods are proposed to improve the usage and coverage level of data based on calculating the distance from GPS point to extend segment domain and using the route map information to predict segment in motion of bus.

The results obtained from the overall study are promising to be a basement to enhance ITS applications such as the multi-criteria route planning in bus network [1], the predicting bus arrival Time based on artificial Neural network and global positioning system data cluster [2].

Acknowledgment

This research is supported by Ho Chi Minh City University of Technology, under grant number T-KHMT-2014-32.

References

[1] Dang Khoa Vo, Tran Vu Pham, Tuong Nguyen Huynh and Van Hoai Tran. “Multi-criteria route planning in bus network” CISIM, 2014

[2] Huynh Nam, Le Viet Hoang Nguyen, Huynh Thanh Quan, Le Thanh Van and Pham Tran Vu. “Predicting Bus Arrival Time Based on Artificial Neural Network and Global Positioning System Data Cluster”, ACOMP, 2014