

USE BIG DATA TO PREDICT SERVICE TIME FOR DELIVERY PLANNING IN CITY LOGISTICS

Liu Shubin¹, Alvin Arulsekaran², Teng Suyan², Derek Ang²

¹Centre of Innovation for Supply Chain Management, Republic Polytechnic, Singapore

²School of Engineering, Republic Polytechnic, Singapore

Introduction

With the flourishing of e-commerce, distribution management also becomes more and more challenging to the logistics and distribution companies. There are more and more jobs with a single parcel at each delivery point. Moreover, customers may require that the delivery to be done within the predetermined delivery time window. The Vehicle Routing Problem with Time Windows (VRPTW), which is at the core of transportation logistics, is commonly seen in real life applications, for example, replenishment of vending machines, delivery of consumer packaged goods to retail/convenience stores and customers, mail collection from post-boxes and mail delivery to households, etc. Extensive studies have been done in the area of VRPTW over the last three decades, with the focus on developing exact and heuristic algorithms for solving VRPTM optimally or near optimally (Laporte 1992).

With an increasing speed of urbanization, urban freight transport system or City Logistics, becomes more and more popular and important to the urban economy. Unlike intercity delivery planning where traveling time plays a significant role, service time (loading/unloading) in city logistics constitutes a high percentage of the overall delivery time. In order to develop an efficient vehicle routing and dispatching plan in city logistics, it is important to have accurate service times. Vleugel and Janic (2004) showed in Amsterdam that approximately 11% of the time is spent driving to the service area, 32% of the time for loading/unloading, and 51% of the time for waiting or driving between customers, and 6% of the time for the driver's lunch time or personal needs. It is clear that these values are city dependent but it also shows that service time plays a critical role in city logistics.

Most of the research work on VRPTW assumes that the service time is known although this may not be true. Accurate service time is essential to forming executable delivery routes, but to most transportation companies, the service time resides in the drivers' brain and is known to the drivers only. Moreover, it is not easy to obtain the service time for new delivery locations which have not been visited before by any drivers.

This paper provides a framework to collect historical data and use the historical big data to retrieve service time. Based on the historical service time, it is able to effectively predict the service time for future delivery jobs. It is noted that we use "delivery" job in this paper but both delivery job and collection job service time can be predicted using the same method.

The main research focus of this paper is to develop a theoretical framework to predict service time based on Global Positioning System (GPS) track data and Electronic Proof of Delivery (EPOD) data. The paper is organized as follows. In Section II, we describe the proposed framework on how to collect GPS and EPOD data as well as how to process the GPS data. Then we present an approach to use GPS track data and EPOD to retrieve service time in Section III. Experimental data analysis is also presented in Section III. Lastly, conclusions and future research directions are summarized in Section IV.

Data Collection and Data Processing

The primary idea here is to analyse historical GPS data and EPOD data to retrieve and predict service time. Based on the GPS data of a delivery man and EPOD data, it is possible to analyse the delivery man's speed (vehicle speed and walking speed) and location for each delivery job and hence retrieve its start and end time. An algorithm was developed to retrieve the total service time starting from the time that a delivery man stops his vehicle to carry out his delivery to the point that he starts his vehicle engine again for the next delivery job. When the transportation companies build up enough delivery historical data, they can predict the service time accurately.

Data Collection

In this study, an Android application was developed to run on Android phone as illustrated in Figures 1a and 1b. The Android application is installed on an Android phone which is carried by the driver or delivery man for the whole delivery trip. The mobile application is used as GPS track logger and is also an EPOD to update job status. Note that the GPS track data will be collected in background automatically once the user logs into the mobile application at the beginning of his delivery trip. The GPS data include: latitude, longitude, timestamp (which is the time and date), speed and accuracy, as shown in Table 1. When a job is finalized, the user will click "Done" in the application to update job status and EPOD data will be captured at the same time. The EPOD data fields include Customer ID, Postal Code, Address and Delivery Time as illustrated in Table 2.



Figure 1a: List of jobs



Figure 1b: Complete a job

LOCAL DATE	LOCAL TIME	LATITUDE	LONGITUDE	SPEED(km/h)	ACCURACY(meter)
14/6/2016	17:52:33	1.3900	103.9103	57	8.4
14/6/2016	17:53:16	1.3855	103.9146	55	9.3
14/6/2016	17:53:40	1.3830	103.9168	45	9.5
14/6/2016	17:54:00	1.3830	103.9168		287.2
14/6/2016	17:54:21	1.3791	103.9232	128	9.5
14/6/2016	17:54:46	1.3784	103.9251	51	9.8
14/6/2016	17:55:10	1.3777	103.9278	51	9.8

Table 1: Sample GPS Track Data

ID	LATITUDE	LONGITUDE	TIMESTAMP	POSTAL CODE	ADDRESS
1323	1.4459	103.7951	14/6/2016 10:17	730758	REMOVED FOR PRIVACY
1324	1.4463	103.7953	14/6/2016 10:23	730758	REMOVED FOR PRIVACY
1325	1.4464	103.7942	14/6/2016 10:29	730759	REMOVED FOR PRIVACY
1326	1.4465	103.7944	14/6/2016 10:32	730757	REMOVED FOR PRIVACY

Table 2: Sample EPOD Data

Data Processing

GPS data acquisition devices have been proven to be useful tools to collect vehicle track data. However, errors may appear in the collected raw GPS data. Common sources of error in GPS data (Adam and Matthew 2012) include sudden signal loss, extraneous or outlying data points, speed drifting and signal white noise. Especially when smart phone is used to collect GPS data, the GPS data accuracy may be affected by indoor environments and the distribution of the nearby cellular towers. As the delivery man will carry the smart phone when he is doing the delivery, the walking speed will also be captured in the GPS track. In this study, we adopted the similar GPS data filtration process proposed by Adam and Matthew (2012). The GPS data filtration process is shown in the following sequence:

1. Remove duplicated records and data with negative differential time
2. Replace outlying data point
3. Fill in speed and timestamp gaps in GPS track data
4. Remove zero-speed signal drift and reset walking speed to zero.

Remove duplicated records and data with negative differential time

This step removes any data points with duplicated timestamp values and data points that have negative or zero differential time. The filter calculates the differential time values for each of the data points and then remove any data points with differential time values less than or equal to zero. We do this step first as it is not possible to estimate missing speed or data points based on differential/integral time information if a duplicated record or negative differential time exists in the GPS track data.

Replace outlying data point

In this step, any erroneous data points, such as single-sample high speed data spikes or missing speed data, are replaced with interpolated data. If the speed in the GPS raw track data falls outside of the range of the pre-set low and high speed limits, the speed will be replaced by the interpolated speed. We derive the interpolated speed by calculating the great-circle distance between two points (Chris 2016), which is the shortest distance over the earth's surface from the neighbouring data point with good accuracy in GPS track data, i.e., accuracy is within 50 meters.

For example, the distance between points (1.3830, 103.9168) and (1.3791, 103.9232) in Table 1 is 833 meters and the differential time is 46 seconds, the derived speed is 65km/hour. Therefore, the missing speed and the high speed spike (128km/hour) in Table 1 are replaced with the derived speed, which is 65km/hour.

Fill in speed and time gaps in GPS track data

The filtration algorithm tries to correct for gaps in the speed-time GPS track caused by any reason, for example, satellite signal is blocked by high rise building, traveling in tunnel or parking vehicle in basement loading bay where satellite signal is lost. The filter examines the timestamp information pulled from the GPS track data and attempts to interpolate any signal gaps which is longer than the time interval based on the sampling rate of the underlying GPS track data. The same interpolation method in replacing outlying data point is used to interpolate speed and time based on the GPS data sampling rate with monotonically increasing time signals.

Remove zero-speed signal drift and reset walking speed to zero

The speed given by GPS device may not be zero when the device is in a stationary state, an effect called "zero speed drift". The GPS device will record a very low speed value, which depends on the sensitivity of the GPS device and satellite signal. To remove the zero-speed signal drift, the filter examines the speed for all data points and the speed is replaced with zero speed value if the raw speed is lower than the pre-set lower speed limit.

As delivery man carries the smart phone when he is walking, the walking speed will be also set to zero in order to detect the smart device's stationary status, which is used to retrieve service time. As the human average walking speed is about 5km/hour, we choose the upper walking speed limit to be 10km/hour in order to consider the variation of walking speed. Therefore any data points with speed less than 10km/hour will be set to zero.

Figure 2 is the raw speed data with zero-speed drift and walking speed. The filtered speed data is shown in Figure 3. As shown in Figures 2 and 3, applying the 10km/hour speed limit makes the track data smoother which is more reflective of the actual delivery operation.

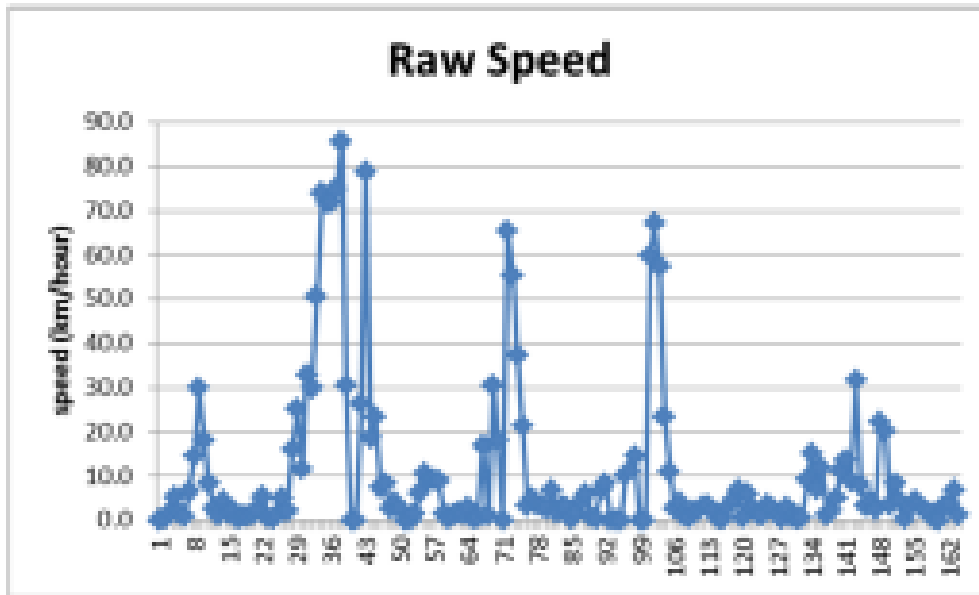


Figure 2: Raw speed in GPS track data

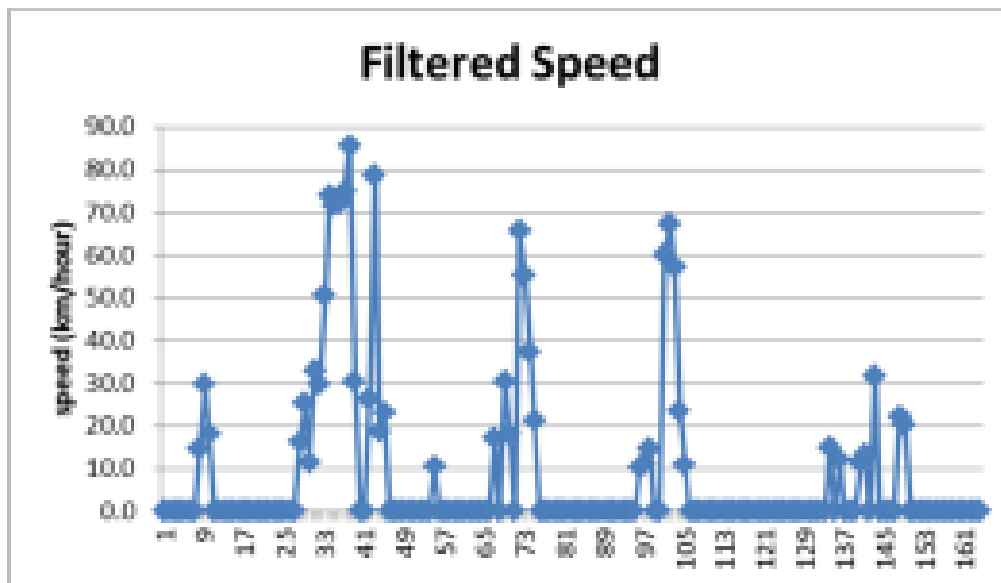


Figure 3: Filtered speed

Approach to Retrieve Service Time

Description of the Proposed Approach

We use both filtered GPS data and EPOD data to retrieve service time of delivery jobs. In the proposed approach, we match the timestamp of EPOD to the GPS track timestamp. Then we trace the vehicle's stationary state and then get the start time and end time of the stationary state based on the zero speed data points in the filtered GPS track data.

The steps to retrieve service time from the filtered GPS data are summarised as follows:

1. Sort the EPOD by non-decreasing order of timestamp
2. Match the EPOD timestamp to GPS track data, if the speed in the GPS track is zero, then find the first and last data points of the vehicle stationary state, the interval of the two timestamps will be the service time. Otherwise, ignore this EPOD as the EPOD data may be invalid.

3. If other EPODs also fall in the time interval obtained in Step 2, then these delivery jobs are a cluster of delivery jobs and share the same service time. Otherwise, go to Step 4.
4. Find the next EPOD which has yet been checked against the GPS track, go to Step 2. If all EPODs have been checked, stop the searching.

Experimental Data

We use the sample EPOD data in Table 2 and the filtered GPS track data shown in Figure 4 as an example to illustrate how the service time is retrieved.

Step 1: Sort the timestamp of EPOD data.

Step 2: The first EPOD data is for job ID 1323 with EPOD time 10:17. From Figure 4, we can get the first zero speed data point timestamp before 10:17, which is 10:15:39, and the timestamp of last zero speed data point in the vehicle stationary state after 10:17 is 10:34:21. Therefore, the service time is from 10:15:39 to 10:34:21, which is 18 minutes 42 seconds.

Step 3: Jobs with ID 1324, 1325 and 1326 are also in the interval of 10:15:39 to 10:34:21. Therefore, the total service time of the four jobs are 18 minutes 42 seconds and the four jobs form a delivery cluster, which means the four jobs are completed by parking the vehicle one time only.

Step 4: Stop as all EPOD data points have been checked against the filtered GPS track. The service time we obtained is shown in Table 3.

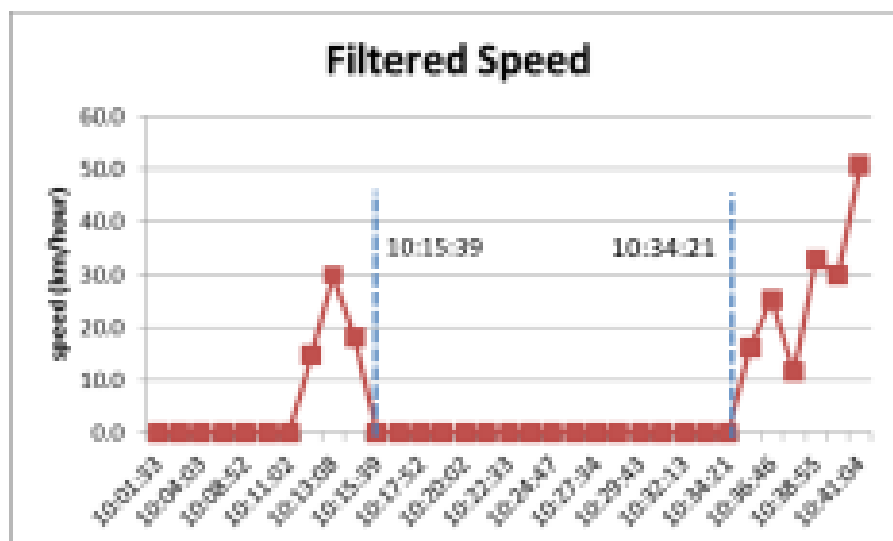


Figure 4: Illustration of service time derived from GPS track data

ID	LATITUDE	LONGITUDE	TIMESTAMP	POSTAL CODE	TOTAL SERVICE TIME
1323	1.4459	103.7951	14/6/2016 10:17	730758	18 minutes 42 seconds
1324	1.4463	103.7953	14/6/2016 10:23	730758	
1325	1.4464	103.7942	14/6/2016 10:29	730759	
1326	1.4465	103.7944	14/6/2016 10:32	730757	

Table 3: Cluster service time obtained from GPS track data

In Singapore, each postal code represents one unique delivery location and hence jobs ID 1323 and ID 1324 are at the same location and share the same "overhead" time, which is the time taken to park the vehicle and walk to/from the building. Based on the postal code data in Table 3, we also find

that jobs with ID 1323&1324, job with ID 1325 and job with ID1326 are at different locations. In this study, we split the cluster service to each individual location. From the delivery cluster, we split the service time of two different locations using the midpoint of the two EPOD timestamps. As illustrated in Figure 5 below, 4 addresses with 3 different locations are visited. As block 758 has 2 delivery jobs, we consider the earliest delivery (job ID 1323) as the first delivery (FD) and the latest delivery (job ID 1324) as the last delivery (LD) for the same delivery location. The same method is used to define any number of delivery jobs in a single block. If the block has only one delivery, it will be both the first and last delivery.

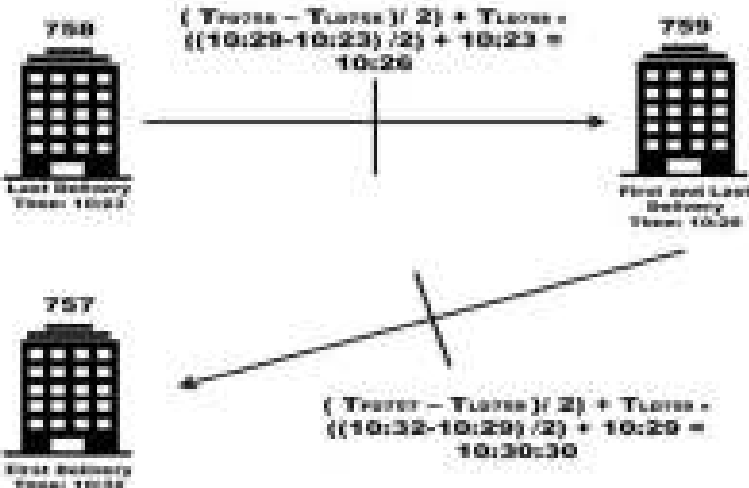


Figure 5: Illustration on how to split the service time

Based on the method described above, the service time is split into three service times based on different delivery locations as shown in Table 4.

ID	LATITUDE	LONGITUDE	TIMESTAMP	POSTAL CODE	SERVICE START and END TIME
1323	1.4458	103.7951	14/6/2016 10:17	730758	10:15:39 to 10:26
1324	1.4463	103.7953	14/6/2016 10:23	730758	
1325	1.4464	103.7942	14/6/2016 10:29	730759	10:25 to 10:30:30
1326	1.4465	103.7944	14/6/2016 10:32	730757	10:30:30 to 10:34:21

Table 4: Split the cluster service time

Conclusions

In this paper, we proposed a framework to retrieve the delivery service time for vehicle routing. Delivery service time can be retrieved from GPS track data and EPOD data. More service time data can be collected when more and more historical data are collected and processed. As the weight and volume of the product delivered by different transportation companies may be very different, the historical service time can only be used by the company itself. Based on the proposed framework, transportation companies is able to predict the service times of delivery jobs based on the historical service time retrieved from their historical GPS track and EPOD data. Therefore, it is possible to develop more efficient and executable delivery routes which are very important for on time delivery service.

In this study, we did not study the relationship between cargo weight, cargo volume, address type, product type, vehicle type and service time. Further work can be done to study the relationship between these factors and service time in order to predict more robust service time. Moreover, the service time may be able to be used cross different companies.

References

- Adam, D. and Matthew, E. (2012). "GPS data filtration method for drive cycle analysis applications", *SAE World Congress*.
- Chris Veness, <http://www.movable-type.co.uk/scripts/latlong.html>
- G. Laporte, G. (1992). "The vehicle routing problem: an overview of exact and approximate algorithms", *European Journal of Operational Research*, Vol. 59, pp. 345–358.
- Vleugel, J. and Janic, M. (2004). "Route choice and the impact of logistic routes", In: Taniguchi, E., Thompson, R. (Eds.), *Logistics Systems for Sustainable Cities*. Elsevier.

Additional Information you

Additional Costs

Students are responsible for paying the tuition, mandatory health insurance and other additional costs.

Master's degree requirements

Students must hold a Bachelor's degree to enroll in the

Admission

Students must apply online and complete the online application process at www.tu-bs.de.

Research Program

Students are encouraged to apply for a research assistant position in one of the research groups of the Business Administration Department. Research assistant positions are available in the following research groups:

Additional Requirements

1. A Bachelor's degree in Business Administration or a related field with a minimum GPA of 2.0.
2. A letter of recommendation from a faculty member at the university of origin.
3. A letter of recommendation from a faculty member at the university of origin.
4. A letter of recommendation from a faculty member at the university of origin.
5. A letter of recommendation from a faculty member at the university of origin.
6. A letter of recommendation from a faculty member at the university of origin.

Tuition Fees and Expenses

Tuition fees and expenses for the first semester are as follows:

For more information contact:

Office of Prospective Students or Business Administration (PhD)
Phone no. +49 331 270-2700, Addressing Building 5
Thomsen Business School, Thomsen University
2 Thomsen Platz, Potsdam, Germany 10625

See www.tu-bs.de for additional fee information.
Email: prospective@tu-bs.de

Website: www.tu-bs.de/en/bsa/bsa-phd-requirements



PhD

OFFICE OF PROSPECTIVE STUDENTS
IN BUSINESS ADMINISTRATION



Master of Engineering (M.Eng.)
Logistics Engineering and Supply Chain Management
 (Bilingual Program)
 物流工程与供应链管理
 (中法合作办学)

Class Location:

Faculty of Engineering
 Zhengzhou University

Duration:

2 years program

Language:

Bi-Lingual (Thai and English)

Qualifications:

Bachelor Degree in Engineering,
 Science, Agriculture,
 or related fields



Application forms and information are available at

The Institute of International Education
 at Zhengzhou University
 (www.zhuizhijiaoyu.com)

For more information

contact the **Logistics Engineering**
 Institute, Engineering Department, Faculty
 of Engineering, Zhengzhou University
 at Tel: 86 371 6388 2114
 Fax: 86 371 6388 2110
 E-mail: li@zhuizhijiaoyu.com
www.zhuizhijiaoyu.com



Now with options
Double Degree with
Ottawa Graduate University
Germany



SCHOOL OF BUSINESS — MASTER OF BUSINESS ADMINISTRATION

WE SET YOU ON THE RIGHT PATH
SO THAT YOU GO EVEN FURTHER.

For more information, please visit
www.unisim.edu.sg/Unisim_MBA



SCHOOL OF BUSINESS

— MASTER IN MANAGEMENT (URBAN TRANSPORTATION)

WE ENABLE YOU TO PLAY A LEADING ROLE
IN POLICY FORMULATION, PLANNING,
MANAGEMENT AND OPERATIONS OF
LAND TRANSPORT SYSTEMS.

For more information, please visit
www.unisim.edu.sg/5812



SCG LOGISTICS

Deliver Your Success



Our Services

- Logistics Solutions Provider
- Domestic Transportation Management
- Ocean Freight Services
- Air Freight Services
- Warehouse Management
- 24/7 Real-time Tracking and Control
- Regional Logistics Coordination and Services

Why SCG Logistics?

- On-time Service
- Strong Relationships
- Proven Network
- 24/7 Real-time Tracking and Control



www.scglogistics.com

