Journal of Engineering and Technology
Vol.4 No.2 July - December 2016

Transportation Modes Detection in Bangkok Using GPS Logger data and GIS Data

Kunnaree Kritiyutanont[1], Masahiko Nagai[2], Apichon Witayangkurn[3]

Remote Sensing and Geographic Information Systems, Asian institute of Technology
P.O. Box 4, Klong Luang, Pathumthani 12120, Thailand
[1]kunnaree.ait@gmail.com
[2]nagaim@iis.u-tokyo.ac.jp
[3]apichon@iis.u-tokyo.ac.jp

***Abstract*— Person trips and transportation mode surveys could be in multiple formats, such as telephone interviews and questionnaires. These data collecting methods rely on manual labeling of data after a survey, and thus, it requires more manpower, time and budget. However, the information technology has introduced advanced data collecting methods such as a mobile phone or a data logger device that can easily record travel time and location data of people. This information plays an essential role in transportation surveying. GPS data can be used to find many features involved in travelling, but those data need to be processed to find transportation modes used before further analysis. The main objective of this study is to detect transportation modes used in Bangkok using GPS logger data. Since the transportation modes in Bangkok are unique and various, there are many problems, such as traffic condition and complexity of the transportation network systems. Therefore, it is not very simple to determine transportation modes. GIS data is used to help detecting transportation modes that have specific routes and stations. Random Forest classifier is used for transportation modes detection. Modes considered in this study are walking, 2-wheel vehicles, 4-wheel vehicles, bus, skytrain, subway and boat. Moreover, activities of people in a week were focused on. Such activities include stationary and modes transferring points. The transportation modes could be automatically detected using our algorithms. This method can be applied for other person trip data collected from a mobile phone that can collect huge number of dataset, and the output data can be used for further analysis in transportation surveying and other related topics.**

***Keywords*— GPS logger, GIS, Transportation modes detection, Random Forest, Bangkok Transportation**

I. INTRODUCTION

In Bangkok, there are many types of transportation systems available, for example, Fig. 1. Traveling in Bangkok could be different from other cities. Bangkok has local transportation modes such as boat and tuk tuk. Many offices and department stores are located in the central of Bangkok. This makes the rate of people travel in and out of Bangkok high in each day.Thus, the problem with traffic congestion has been around for a while and very hard to deal with since the government agencies, that deal with this problem, still lack the information about "people movement". This information can be used to analyze to improve or adjust the transportation systems in Bangkok to satisfy the needs of people.

To collect information on person trips and transportation modes, it can be done using several ways. Fortunately, theadvanced data collecting devices (e.g. mobile phone or data logger device) have been introduced. These devices can automatically record travel time and location data of people without them having to record the data manually. GPS data can be used to find features involved in traveling, but the data needs to be processed prior to further analysis. The information on the transportation modes used by people can reflect the lifestyles of people as well. However, it is not feasible to have every user manually tagging the corresponding transportation modes to their GPS tracks. Two obvious reasons are:

1) Users do not directly achieve benefits from labeling their trips, and, 2) It is not really convenient for people to record the exact time that they change modes of transportation [1]. From pre-field experiment, the people in the study group do not feel comfortable to use their phones to record data, but GPS logger can record data for a longer period of time than mobile phones. GPS logger can last up to 4-5 days after a full charging.

Since GPS logger devices cannot detect the transportation modes used automatically, it needs a calculation method to find transportation modes. Moreover, GIS data can be used to improve accuracy of detection algorithm for modes in Bangkok that have specific routes and stations.

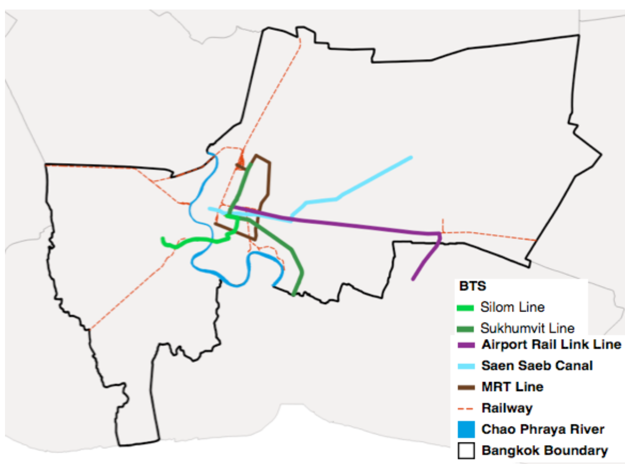

Fig. 1 Transportation networks in Bangkok

II. OBJECTIVES

The main objectives of this work are based on 3 followings:

• The first objective is to find suitable procedure for data collection and manipulation of GPS logger, questionnaire and GIS data. Since this work requires data from many sources, the suitable way, to collect and manipulate data, has to be designed in order to reduce the time used and further work.

• The second objective is to detect transportation modes in Bangkok using GPS logger and GIS data. Classification features and algorithms are adapted to match with transportation in Bangkok.

• The third objective is to analyze the detected transportation modes using grid and overlay technique, and create moving objects visualization.

III. RELATED WORKS

The approach to automatically identify users' transportation modes in [8] has three parts which are "change point-based segmentation method", "inference model", and "graph-based post-processing algorithm". The GPS trajectories are partitioned into segments of different transportation modes. Then, a set of features, such as direction change rate, velocity change rate, and stop rate, is identified and input to the inference model to classify the segments of different modes. These features are not affected by differing traffic conditions. Finally, the graph- based post-processing algorithm is conducted. This algorithm considers the commonsense constraints of the real world and users' behaviors based on locations in a probabilistic manner. From the results, at least 71 % prediction accuracy is achieved. In [5], a user's transportation modes are identified based on the GPS sensor on the user's mobile phone and transportation network data. The transportation network data includes real time bus locations, spatial rail and spatial bus stop data. By having the transportation network data, the motorized transportation modes can be identified with much higher accuracy than without having the transportation network data. From this paper, five different inference models including Bayesian Net, Decision Tree, Random Forest, Naïve Bayesian and Multilayer Perceptron were used to test. Among all five mentioned inference models, Random Forest performed best with over 93% precision and recall accuracy in the experiments conducted. In [7], an approach to reconstruct user's trip information from low-data-rate GPS data from a mobile phone has been proposed. There are four steps in the approach, which are "stay point extraction with outlier detection and removal", "trip segmentation based on change point detection", "transportation mode extraction using inference model" and "segment merging". During the "stay point extraction", outlier detection method is used to detect outliers and combine them as stay points. For "trip segmentation based on change point detection", some extra features, such as velocity change rate and point in train, are considered to assist the segmentation of non-walk segments that contain multiple modes. Spatial features including spatial train network and spatial road network are employed to improve the transportation modes inferring. After adding the spatial train network and spatial road network, the results achieved show 87.8% and 84% accuracy for precision and recall, respectively. Random Forest classifier with Bootstrap Aggregating is used for classifying transportation modes. From the results, this approach could reach over 87.8% accuracy for inferring transportation modes and 97.76% for just train mode alone.

IV. MODES OF TRANSPORTATION IN BANGKOK

There are several transportation systems currently operating in the city of Bangkok (see more in Fig. 2). Although Bangkok's canals had served as the major transportation mode in the past, the land traffic has surpassed them in terms of importance. There are also older public transportation systems still operating in the city. Such older transportation systems include extensive bus networks and boat services on the Chaophraya River and Saen Saeb canal. Unfortunately, in the present days, Bangkok's public transportation systems cannot fully satisfy the lifestyles of the majority of its people in terms of quality and safety. Many people in Bangkok would still prefer to use their private vehicles over other public transportation modes. This is probably one of the very main reasons that limits the income of the public transportation systems and makes it harder to improve the service [6].

Fig. 2 The examples of modes of transportation in Bangkok

The algorithm and classification features used in previous works are different. It depends on their data rate of GPS recording, device used, characteristics of transportation modes to detect and transport network data available. For example, [5] used real time bus location that, in Bangkok, is not available. For Bangkok, there are traffic conditions and various kinds of modes that can affect the results.

Considering some features, such as velocity and acceleration, might not be enough to infer the transportation modes in Bangkok since traffic condition could get really bad that inferring each transportation modes is almost impossible to be accurate. Modes, with the same characteristics and no available data to be used for detection, should be combined into the same group. Using GIS data to help infer the transportation modes make the results more accurate. Moreover, from previous reviewed researches, using Random Forest classifier to infer the transportation modes produces more accurate results than using other classifiers. Segment merging technique also helps reduce uncertain segments and traffic problems related in the detection step that affect accuracy when separating walk and other modes. Thus, it is used in this work for transportation modes detection step.

V. METHODOLOGY

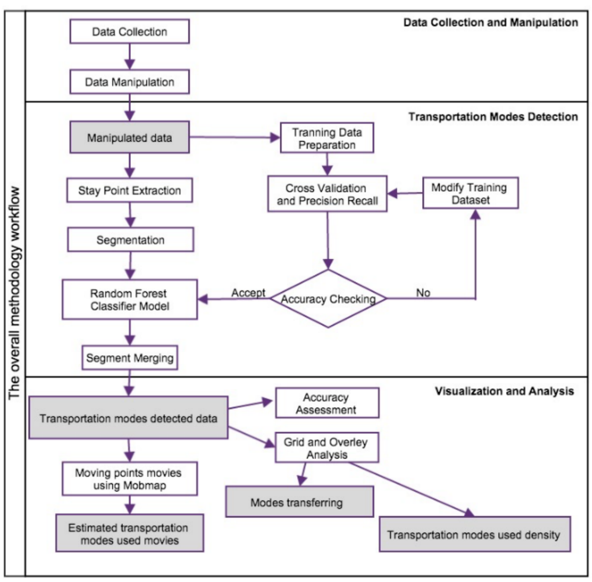

Fig. 3 Overall workflow of methodology

For the methodology (Fig. 3), it is divided into three main parts:

1) Dataset Collection and Manipulation: data is collected from GPS logger, questionnaire survey papers and GIS data. After all the needed data has been collected, it is processed to be used in the other parts.

2) Transportation Modes Detection: the detection methods include stay point extraction, walk and non-walk segmentation, and transportation modes detection using Random Forest classifier and training data from the study group of volunteers.

3) Visualization and Analysis: after the transportation modes used have been determined, the information is then visualized and analyzed to find mode-transferring points and main transportation modes used in Bangkok.

VI. DATASET COLLECTION AND MANIPULATION

For person trip data, there are 80 volunteers in total from 4 universities. Each university has 20 volunteers. The volunteers who tended to use various types of transportation modes were preferred so the variety of data on the transportation modes could be achieved. The volunteers came from 4 universities, including Kasetsart University, Chulalongkorn University, Srinakharinwirot University and King's Mongkut Institute of Technology Ladkrabang. All of these universities are located in Bangkok but in different areas. The local transportation modes for the students from these universities are different.

A. GPS Logger Data

The device that is used to collect data is "GPS logger igotu GT-600" from Mobile Action, shown in Fig. 4. This device can set the time to record, has motion detection, and is water resistant. This device is selected because it is easy to carry around. From the field experiment, it could operate for about a week without charging while recording data every 5 seconds. This experiment uses 5-second time interval to record the data. Motion detection on the devices has been turned on to reduce number of records when people have no movement, and the buttons on the devices have been locked to prevent anyone from causing any problem to the devices. The period for collecting data from each volunteer is one week.

Fig. 4 igotu GT-600 GPS logger devices

B. Questionnaire Survey Papers

The volunteers recorded their person trip in the given questionnaire survey papers while they are having the GPS loggers. The complete ones must complete details of the starting and ending time of each trip, modes used, origin and destination. The transportation modes used and the corresponding times are recorded. This information is used to improve the accuracy of inferring and create training dataset.

C. GIS Data

Transportation network data is achieved from the related government agencies and downloaded from the BMA GIS Center's website, which provides free GIS data. Digitizing is performed for GIS data that are not available (e.g. MRT entrances, platform area of the sky train and bus stop). In Fig. 5-6, the examples of the transportation modes with specific routes and stations are shown.

D. GPS Logger Data Manipulation

After importing data from the GPS logger devices, it is noticed that each volunteer's data consists of many parts. So, they need to be combined. Java and PostgreSQL database is used to develop a program to combine the parts, and add or delete as desired. The example of GPS data processed is shown in Fig. 7, only timestamp and location data are extracted.

Fig. 5 Entrances of Lat Phrao MRT station

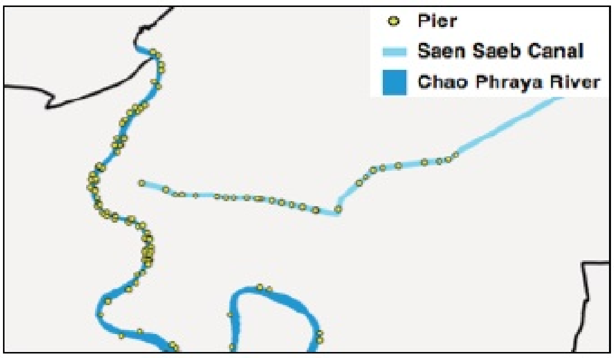

Fig. 6 Routes and piers of boat services in Bangkok

```
userid,sampleid,deviceno,time_stamp,latitude,longitude,altitude
cu01,1,120055,2014-09-30 00:00:01,13.755944,100.533295,50.97
cu01,1,120055,2014-09-30 00:00:06,13.755961,100.533318,50.48
cu01,1,120055,2014-09-30 00:00:11,13.756051,100.533394,49.9
cu01,1,120055,2014-09-30 00:00:16,13.756098,100.533447,53.54
```

Fig. 7 GPS data included ID of users and new timestamp format

After the calculation and inspection, the error sources are duplicated records and jumping GPS points. From the calculated data below (Fig. 8), there are some points that should not exist because they would require velocity, acceleration or distance that a vehicle cannot achieve. Those points are detected as noise data (Fig. 9), then noise-removing algorithm is applied in this step.

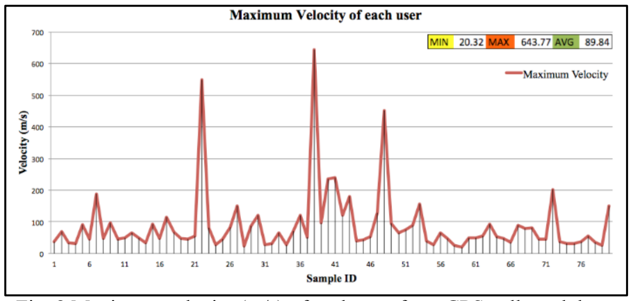

Fig. 8 Maximum velocity (m/s) of each user from GPS collected data

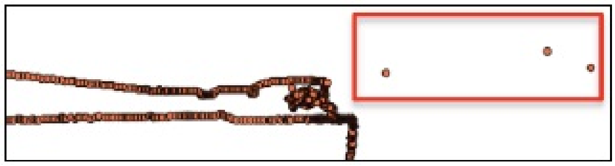

Fig. 9 Jumping GPS points (in the box)

E. Questionnaire Data Manipulation

The questionnaire survey papers of 1- week person trip data are imported into CSV format. One record of trip data includes starting time, ending time, origin, destination and modes used. These data are prepared and collected in database to create training data in further step.

F. GIS Data Manipulation

In this part, GIS data of routes and stations are used to find suitable buffering distance of each data to create buffering area. Since GPS data might contain errors, spatial features are used to help increase the accuracy of finding GPS points on the particular routes and stations. Moreover, all GIS data are changed projection into EPSG: 4326 same as GPS data.

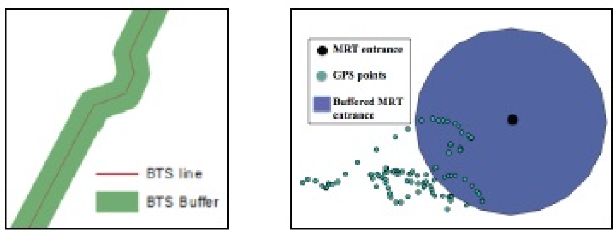

Fig. 10 Bufferd GIS data

VII. TRANSPORTATION MODES DETECTION

A. Stay Point Extraction

The area with the cluster of GPS data (Fig. 11) is determined to identify if the user "stay for a while". The rest will be "move". Only moving part will be further analyzed. The method that is used to identify them is stay point extraction technique [7]. Stay points are extracted by using "minimum time" (tmin) and "maximum distance" (dmax) as criteria. Values selected to be thresholds in this work are tmin=15 minutes and dmax = 200 meters. So, if there are a GPS points hit thresholds (tmin>=15 and dmax<= 200), those points are detected as stay, then only move segments are further analyzed.

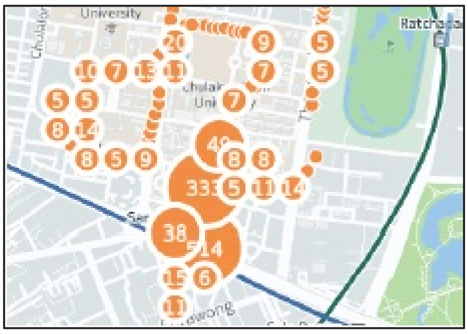

Fig. 11 Clustering of GPS points

B. Walk and Non-Walk Segmentation

Walk segments are analyzed to separate walk and non-walk segments. In this part, velocity (v) and acceleration (a) are used as 2 criteria. When v=2 m/s and a=0.6 m/s, It got high recall with the best precision. Then, only non-walk segment are used in the next step for transportation modes detection. Moreover, to avoid GPS signal losing and missing segments when people getting in sky train or subway, walk segments are used to perform within function with sky train platform and subway entrances data. If there are time gaps between two walk segments hitting thresholds and those walk segments are along with sky train platform or subway entrances, the velocities, at the ending point of first segment and starting point of current segment, are checked again to avoid some cases, such as people walking on BTS-skywalk or just cross the road using MRT entrances. If those segments are accepted, they are merged into one non-walk segment.

C. Classification Features

To detect transportation modes in Bangkok, suitable classification features are required. Considering just some features, such as basic features, might not be enough to infer the transportation modes in Bangkok. Using advanced features, such as Stop Rate and Heading Change Rate (Fig. 12) and GIS data, to help infer the transportation modes will make the results more accurate.

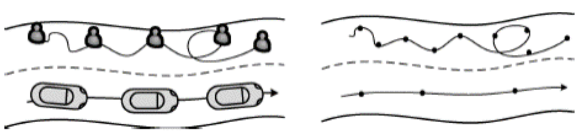

Fig. 12 Heading change between walking and driving in the same distance [8]

Since some of the transportation systems in Bangkok have similar routes, such as a boat route with a bridge above (Fig.13). If any movement on the water area is considered as a boat, the accuracy achieved could be very low since there are many bridges for land vehicles to cross the rivers. Thus, Stop Rate at Piers of a segment is used to separate boat and others. The vehicles travel on the roads under the BTS railway could also be considered as BTS mistakenly (Fig. 14). So, this work introduced "the percentage of stop rate at stations" as new spatial features. Sky train stops only at its stations (Fig. 15), but vehicles travel along the sky train lines might stop at other parts such as intersection. This characteristic can be used to identify the modes.

There are differences in the driving behaviors between motorbike users and car users. For example, motorbike riders always zip in and zip out through traffic around Bangkok so the angle of heading of that segment can change more than using cars [8]. For a bus, people who taking a bus are likely to stop more times than driving, so stop rate can be another possible feature, but during peak hour, this feature cannot be used. Thus, the possible feature that can extract bus from another is stop rate at bus stop of that segment.

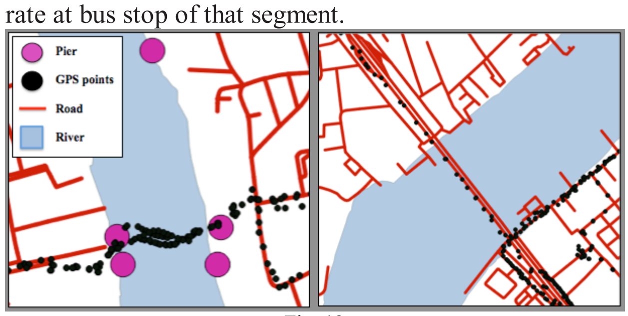

Fig. 13
(Left) Boat segments on Chao Phraya River
(Right) Vehicles on the road segments

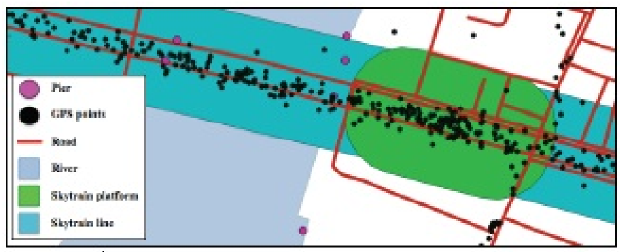

Fig. 14 Complexity of transportation networks between boat, sky train and modes on the road.

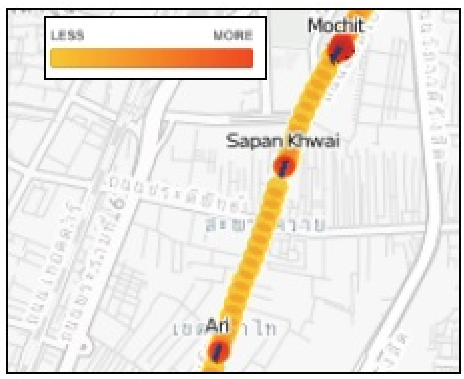

Fig. 15 Cluster of GPS points along the BTS line

The classification features used in this work are shown in Table 1. Features ranking is also performed to help in features selection.

Below are the examples of equation to calculate classification features.

$$HCR = P(hcr)/Dist \qquad (1)$$

Where P(hcr) is the number of GPS points in a segment which a user changes his/her heading direction exceed a certain threshold.

$$SR = P(sr)/Dist \quad (2)$$

Where P(sr) is the number of GPS points in a segment which velocity is smaller than a certain threshold.

$$PiT = P(pit)/\text{Total point in a segment} \quad (3)$$

Where P(pit) is the number of GPS points in a segmen which belong in a particular sky train line.

$$PiST = P(pist)/P(sr) \quad (4)$$

Where P(pist) is the number of P(sr) within sky train platforms.

TABLE 1
CLASSIFICAITON FEATURES

Basic Features	
Dist	Total distance of a segment
MaxV	The maximal velocity of a segment
MaxAc	The maximal acceleration of a segment
AvgV	The average velocity of a segment
Time	Travel time of a segment
Point	Total point of a segment
Advanced Features	
SR	Stop rate of a segment
HCR	Heading Change Rate of a segment
Spatial Features	
PiT	Percentage of point in sky train line
PiW	Percentage of point in River and Canal
PiSW	Percentage of point in subway entrances
PiST	Percentage of stop rate in sky train platforms
PiP	Percentage of stop rate at piers
PiB	Percentage of stop rate at bus stops

D. Transportation Modes Classification Model

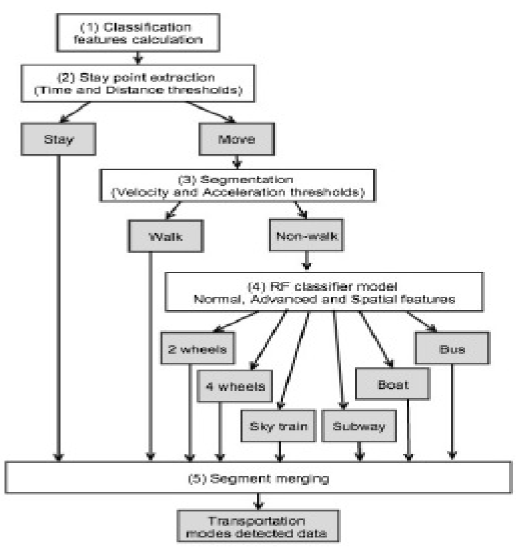

Fig. 16 Workflow of modes detection

Random Forest is used for transportation modes detection.Dataset and features combination that obtain the best result of precision and recall are selected to be training data. All remaining GPS data are calculated to find classification features based on segment to be used as test data in Random Forest model.

For the training data, the data from questionnaire survey papers are used to find threshold and create training data for learning state of Random Forest model. In this work, out of all the data, about 30% of them are acceptable and used as training data.

E. Segment Merging

In this part, the probability of the transportation is analyzed. For example, in a short distance, it is unlikely that a user will use many transportation modes. This part of data needs to be corrected. Moreover, a walk segment is analyzed. If it is in between two transportation modes (Fig. 17), it is a possible walk segment. If it is in between same transportation mode, it should be merged to that mode because the velocity of the mode could be decreased due to the traffic condition.

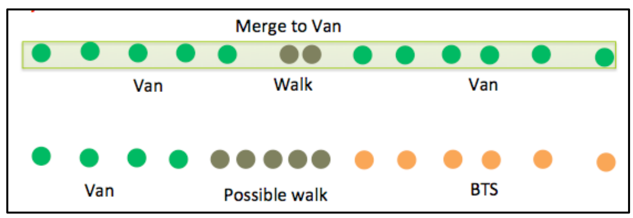

Fig. 17 Certain and uncertain walk segments

F. Performance Evaluation

Precision and recall are the measures used in evaluating model. K-fold Cross Validation is a method of assessing the accuracy and validity of the model.

VIII. VISUALIZATION AND ANALYSIS

To visualize the results, labeling data by modes in integer makes it easier to visualize the data on GIS softwares and other applications (e.g. QGIS and Mobmap). It also reduces the size of the data. Mobmap application is used to create moving points visualization. Detected mode data is used as the input file (Fig. 18). Each mode is identified by different colors. Then, it is generated to moving GPS points, separated by user's ID, sorted by timestamp and colors varying when its modes are changed.

For Grid and Overlay Analysis, detected transportation modes data are used to perform spatial joining with grid covered Bangkok areas.

sampleid	latitude	longitude	new_timestamp	modes	mode_int	transfer
1	13.751527	100.530952	2015-06-02 08:34:45	walk	1	0
1	13.751505	100.530991	2015-06-02 08:34:50	walk	1	0
1	13.751491	100.531067	2015-06-02 08:34:55	walk	1	0
1	13.749519	100.531212	2015-06-02 08:35:00	bus	4	0
1	13.74901	100.531082	2015-06-02 08:35:05	bus	4	1
1	13.748729	100.531036	2015-06-02 08:35:10	bus	4	0
1	13.748468	100.530945	2015-06-02 08:35:15	bus	4	0

Fig. 18 Examples of modes detected data and its new modes labeling

From the data, information such as "modes transferring point" which is a point that people change their transportation mode can be found. In Fig. 18, the record which value=1 in transfer column, it is collected in modes transferring points. In addition, each grid might contain several modes used, so major modes can be found by finding the highest percentage of modes used. The data in this work is spatio-temporal GPS data, so the structures of grid are created based on hour of a day (Fig. 19) for further analysis and visualization.

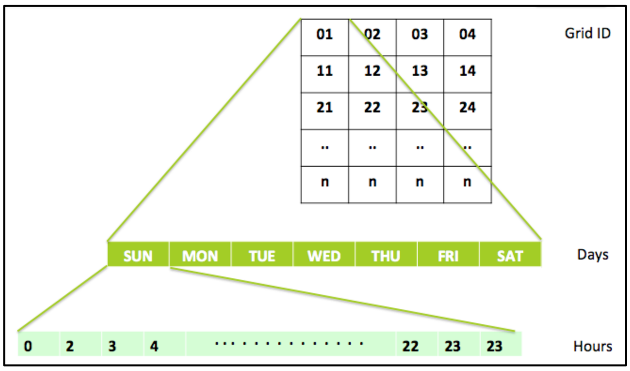

Fig. 19 Grid structures

For the implementation, PostgreSQL with PostGIS extension is used as database to collect and manage spatial and non-spatial data. Java language is used to develop program for calculation and modes detection. For Random Forest classifier, Java-ML library is used to develop the model. Finally, QGIS, Mobmap and CartoDB are used for visualization. Moreover, those are available for working in a cross platform.

IX. RESULTS AND DISCUSSION

Comparing to questionnaire surveys, GPS devices can automatically record time and location of person trip, whereas using questionnaire surveys, people must report their trips manually. In Fig. 20, the information about active periods from GPS is shown. However, without processing and analysis, questionnaire data can provide more information than raw GPS data.

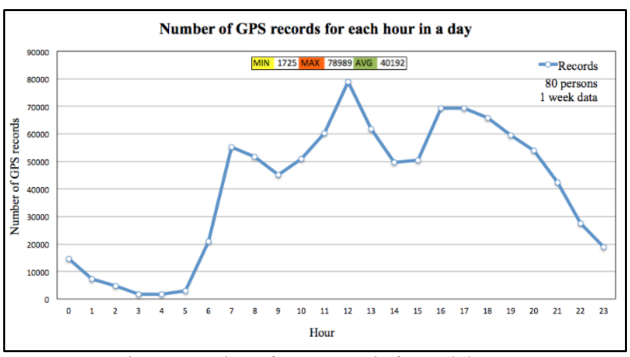

Fig. 20 Number of GPS records for each hour

After manipulation and stay point extraction, the process of GPS noise removing makes the GPS data usable for analysis. From the results shown in Fig. 21, starting and ending time of each stay segment are defined. The collections of stay points are used to calculate for its centroid location. This information can be used for further analysis about mining activities of people.

There are various modes on the road in Bangkok such as private car, bus, taxi, van, motorbike and etc. In this work, modes on the road detected are divided into 2-wheel vehicles, 4-wheel vehicles and bus, due to same characteristics. 2 wheels; include bike and motorbike. 4 wheels; include car, taxi, van, tuk tuk, minibus and etc.

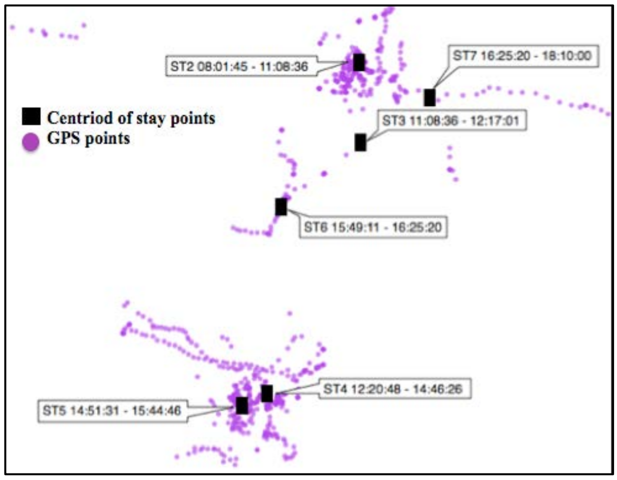

Fig. 21 Examples of stay points detected

For transportation modes detection, results are shown in the figures below. In Fig. 22, it shows modes detected of 1-day person trip data after segment merging is applied. GPS points are separated into several segments and categorized by modes used. For sky train and bus, percentages obtained are lower than other modes due to traffic conditions and complex transportation networks in Bangkok. The overall model precision is 87.17 % and 87.83 % for recall, by using spatial features together (Table 2).

Fig. 22 The example of modes detected data

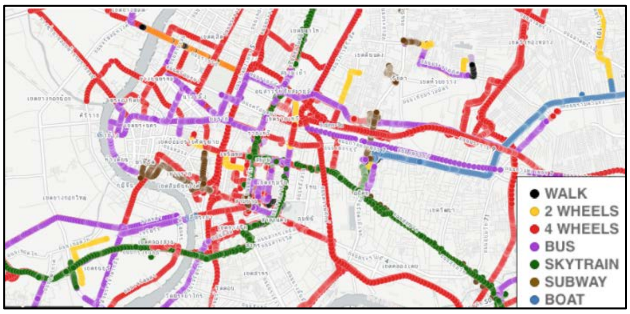

Fig. 23 Modes detected data of several person trips3

For the analysis part, transportation modes detected data and vector grids are processed to find major modes used. In Fig. 24, the results obtained show that boat is only found in Saen Saeb Canal and some parts of Chao Phraya River around the inner Bangkok, which is correct. For modes transferring points (Fig. 25), grids that have rapid transit show a lot of modes transferring points. The darker color grids cover areas such as Mo Chit, Sam Yan, Lat Phrao, Victory Monument and Phaya Thai, that are commercial zone and people can transfer to many other modes.

TABLE 2
COMPARISON THE PERFORMANCES OF TRANSPORTATION MODES DETECTION MODELS

	Basic + Advanced Features			Basic + Advanced + Spatial Features		
Mode	Accuracy	Precision	Recall	Accuracy	Precision	Recall
2 Wheels	90	70	92	96	86	94
4 Wheels	82	67	71	92	86	86
Bus	74	60	40	80	71	70
Skytrain	90	75	43	92	80	77
Subway	92	54	20	100	100	100
Boat	90	73	92	100	100	100
Overall	**86.33**	**66.50**	**59.67**	**93.33**	**87.17**	**87.83**

Fig. 24 The example of grid data (100 m.) that contain boat as the major mode

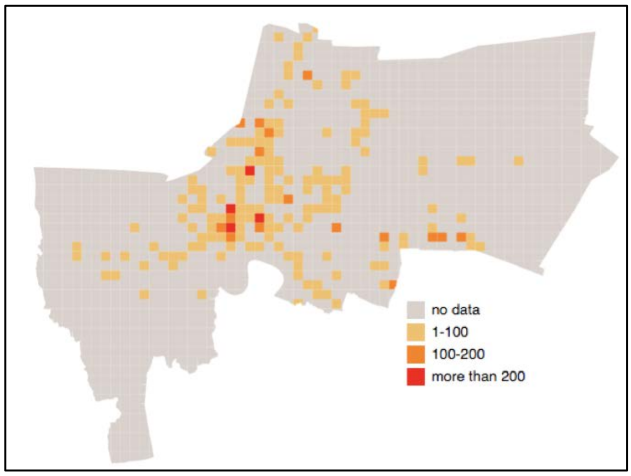

Fig. 25 Modes transferring points in Bangkok

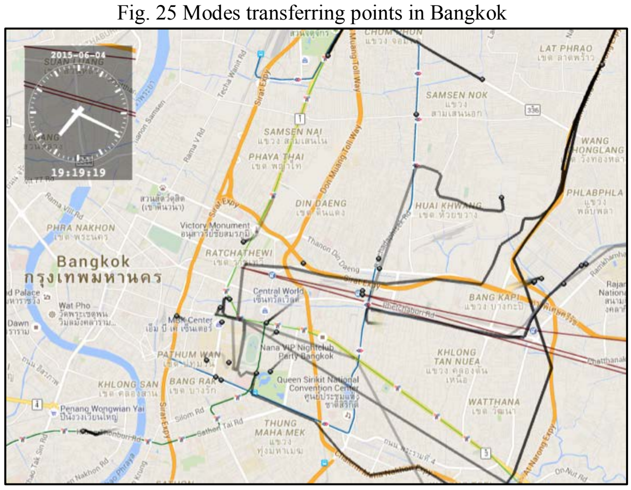

Fig. 26 People's movement from raw GPS data

From Fig. 26, it shows screen capture of user's movement's GPS points generated from Mob map. From Fig. 27, it shows people flow in different colors imply different modes of transportation from transportation modes detected data. This kind of visualization is useful for studying about mobility of people.

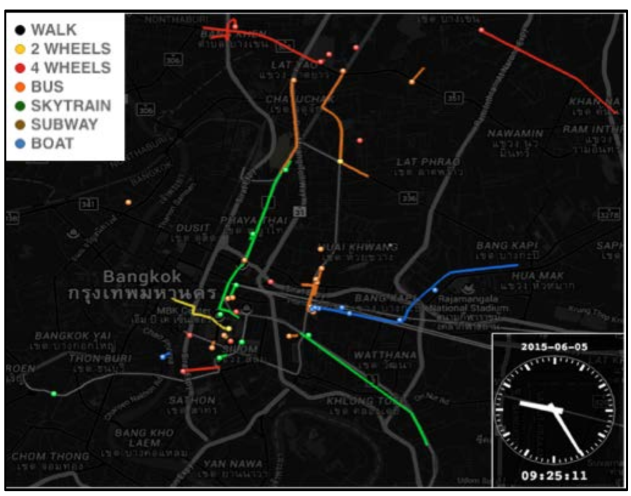

Fig. 27 Estimated transportation modes used

X. CONCLUSIONS AND FUTURE WORKS

A. Conclusions

For the transportation modes detection, conducting the "stay point extraction" help reduce the number of data to be analyzed further. Random Forest classifier and classification features can detect transportation modes used in Bangkok. Using spatial features along with other features gives higher precision and recall than using other features alone. The process of segment merging helps reduce uncertain segments, such as the difficulty determining between walking and facing traffic congestion on the road, to make the person trip more complete. The raw dataset in this work are only timestamp and location that are minimum requirement of GPS log data, so the algorithms and features in this work can be applied for GPS data from other sources such as a mobile phone.

Other interesting features that could help detect other modes (e.g. vans or taxis) are parking area, van terminals, locations of getting on/off a car and probability of the modes transferring. Moreover, to detect modes like buses and motorcycles, traffic information, intersection locations, and real time bus locations, can help increase the accuracy of the detection, if those data are available.

Data collected from the 80 volunteers seem to be insufficient for analysis when comparing to how large the population in Bangkok is. The volunteers's travel data are limited to just around the universities and their houses, shown in Fig. 28.

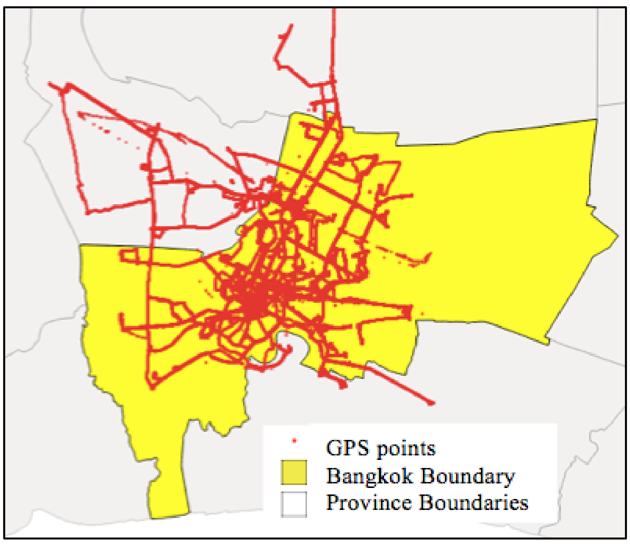

Fig. 28 Distribution of the person trips data collected from the GPS logger

These data obtained could be useful for many other works, such as observing people's traveling demand, planning the expansion of the transportation network, studying people movement when certain events occur, and etc. If this information has been collected and managed, the planning of urban and transportation systems could be easier and more suitable.

B. Future Works

1) To collect GPS data to analyze the transportation modes used by people in Bangkok, there should be larger group of people to be studied. The data might be collected from mobile phones.

2) The process of noise removing cannot remove the entire noise data. More advanced method might have to be adopted.

3) To detect some local transportation modes (e.g. bus, private car and taxi), other classification features, more precise GIS data and more complicated techniques (e.g. using traffic data or pattern of the taxi drivers) should be used.

4) Some data is lost due to the loss of GPS signal and motion detection function; the lost part of the data should be interpolated to fill out the missing part.

REFERENCES

[1] E.-H. Chung and A. Shalaby, "A Trip Reconstruction Tool for GPS-based Personal Travel Surveys," *Transportation Planning and Technology*, vol. 28, no. 5, pp. 381–401, Oct. 2005.

[2] E. D. Kaplan and C. Hegarty, *Understanding GPS: Principles and Applications*, 2nd ed. Norwood, Massachusetts: Artech House, 2006.

[3] L. Liao, D. J. Patterson, D. Fox, and H. Kautz, "Learning and inferring transportation routines,"*Artificial Intelligence*, vol. 171, no. 5, pp. 311–331, Apr. 2007.

[4] S. Reddy, M. Mun, J. Burke, D. Estrin, M. Hansen, and M. Srivastava, "Using Mobile Phones to Determine Transportation Modes," *ACM Trans. Sen. Netw.*, vol. 6, no. 2, pp. 13:1–13:27, Mar. 2010.

[5] L. Stenneth, O. Wolfson, P. S. Yu, and B. Xu, "Transportation Mode Detection Using Mobile Phones and GIS Information," in *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, New York, NY, USA, 2011, pp. 54–63.

[6] "Transport Data and Model integrated with Multimodal Transport and Logistics," *The Office of Transportation and Traffic Policy and Planning, Ministration of Transport*, 2011. [Online]. Available: http://www.otp.go.th.

[7] E.-H. Chung and A. Shalaby, "A Trip Reconstruction Tool for GPS-based Personal Travel Surveys," *Transportation Planning and Technology*, vol. 28, no. 5, pp. 381–401, Oct. 2005.

[8] Y. Zheng, Y. Chen, Q. Li, X. Xie, and W.-Y. Ma, "Understanding transportation modes based on GPS data for Web applications," *ACM Transaction on the Web*, vol. 4, Jan. 2010.

[9] Y. Zheng, Q. Li, Y. Chen, X. Xie, and W.-Y. Ma, "Understanding mobility based on GPS data," *Microsoft Research*, Sep. 2008.