Challenges in Implementing Crowdsourcing on Automatic Real-Time Transit Tracking System

Phuthipong Bovornkeeratiroj, Kulit Na Nakorn, Kultida Rojviboonchai*
Department of Computer Engineering
Faculty of Engineering,
Chulalongkorn University
Bangkok, Thailand
Phuthipong.B@student.chula.ac.th, Kulit.n@student.chula.ac.th, Kultida.r@chula.ac.th

***Abstract*—Implementing Crowdsourcing in arrival time prediction system for public transportation lets people share their pieces of information that they have, such as the transport location or situation, and combine the information together to get the complete picture of the interesting public transportation. However, the system that requires human to input the information has critical problems, which are the needed of the user effort and the fraud by false reporting. So the automatic system has been proposed to solve those problems by detecting necessary information automatically. But still there are some new challenges coming with it. In this paper, we aim to demonstrate those new challenges in automatic crowdsourcing system, which are Performance, Consistency, and Continuity of Tracking and also introduce some possible solutions for those challenges. We tested our solutions by simulating the problem situations with our boat arrival time prediction system, which is an automatic arrival time prediction system for Chao Phraya Express Boat in Bangkok, Thailand.**

Keywords—crowdsourcing; transit tracking; mobile phones; sensors; participatory sensing; boat;

I. Introduction

Crowdsourcing is a method to get certain information, usually in real-time, by gathering many pieces of information from many people. With an era of smartphone, crowdsourcing becomes more popular since it is very easy to share their information via their own mobile phones. Transit-tracking is one of the hot topics that applies crowdsourcing to create a system that lets passengers share the location and information of their transit vehicles with other passengers. For example, if every passenger who is on the bus shares his/her bus location and bus information to the crowdsourcing system, other passengers who are waiting for the bus can see where is their waiting bus and how long they have to wait. The benefit of crowdsourcing in transit tracking is that the system is cheap and independent. It is cheap because no infrastructure, such as GPS tracking on the vehicles, is needed. And it is independent because it does not rely on any public transportation company. To let the user manually input the transit information into their mobile phone every time can discourage the user to keep sharing their information. The user can also deliver the false-report to the system, which can decrease the system accuracy and trust from the users.

In order to solve the above-mentioned issues, another type of crowdsourcing for real-time transit tracking is introduced which is an automatic system. This system does not require any manual input from the user. It can automatically detect necessary information, such as the user is on the vehicle or not and what vehicle it is, automatically in background process and share the information back to the community by itself. Though, the problems of manual input system are solved, it introduces new challenges that need to be considered.

Naveen Nandan, et al [2] identifies common challenges in crowdsourcing transit tracking system. There are 7 challenges, which are: (1) Device Hardware, which concerns about battery life and battery consumption. (2) Network Coverage, the network connection issue can cause the mobile phone to send less information than the system required. (3) Costs, smartphone is still expensive but crowdsourcing need a lot of people for a better performance. (4) Data Availability and Quality, some data need to be prepared before the system can operate, such as map or cell and Wi-Fi for each location, and how to assure that the received information is accurate and correct. (5) Trust, user privacy and source of information should be an anonymous record. (6) Motivation, find a way to keep users sharing their information for a long time. (7) Usability, effort that user need to spend on the system should be at minimum.

Though, many challenges of typical crowdsourcing system have been identified but for an automatic system, there are still more challenges: (1) Tracking Performance, the system must find the best result with the least battery consumption. (2) Tracking Consistency, what if the user leaves or a new user gets in the vehicle. The system should be able to maintain the transit tracking with the best possible result. (3) Tracking Continuity, one of the disadvantages of the automatic detection system comparing to the manual is that some detections cannot instantly be detected, they needs some time to collect certain amount of data. So the system should be able to continue those detections and pass over the detection in any situations.

To the best of our knowledge, there is no work that tries to identify and solve the challenges in the automatic detection system yet. But there are some works that tries to solve manual input system problem such as You Better Be Honest [3], which tackles the false-report problem. In this paper, we focus on the automatic system's challenges by simulating the situations

*Corresponding author

978-1-4673-6820-9/15/$31.00 ©2015 IEEE

using our Boat Arrival Time Prediction System [4], which is an automatic detection system, and demonstrate some possible solutions to those challenges.

The rest of this paper is structured as follows. In section II, we review and summarize related works, including manual input and automatic detection transit tracking system. In section III, system's challenges and solutions are described in detail. Then, in section IV, the solutions to the challenges evaluations are shown. Finally, section V is the conclusion.

II. Related work

There are many researches and existing applications that implemented the transit-tracking system that apply crowdsourcing technique. Some require passengers to manually activate the application and input some necessary data, and some can work without any effort from the passengers. In this section, we review both manual system and automatic system.

ContriSense:Bus [5] allows passengers to share their bus location with other passengers who use the system via their smartphones. After the passengers activate the application, it will keep collecting user location and send to the server for report and predict the arrival time of the bus. The passenger can also report how crowded the bus is. Tiramitsu [6] shares a similar purpose with the ContriSense:Bus but the passengers can report more information such as riding experiences, and disabilities supported. Moovit [1] is also similar to previous researches but it is already deployed and supports many major cities in US and some major cities in the world.

Cooperative Transit Tracking using Smart-phone [7] is a crowd-sourced transit tracking which does not require any input. It utilized sensors in mobile phone such as GPS, Wi-Fi, and accelerometer to automatically detect that the user is already on the transit vehicle, which can be either train or bus. How long to wait? [8] is a mobile phone based participatory sensing for predicting bus arrival time. It is also an automatic system, which has to run on background all the time. But it does not use GPS so the mobile phone consumes significantly less amount of battery power.

The system that we use for demonstrate the challenges and solutions, Boat Arrival Time Prediction System [4], our proposed automatic real-time boat tracking system. The goal of the system is to predict the boat arrival time so that the boat passenger can know the estimated waiting time by just using passengers' mobile phones. Before the arrival time can be computed, the system must complete 4 detections, which are On-Boat Detection, Boat-Stop Detection, Boat-Direction Detection, and Boat Classification. The On-Boat Detection is to detect that the user is embarking the boat and disembarking the boat. After the user is on the boat, the Boat-Stop detection is activated to detect when the boat stops at any pier. The Boat-Direction Detection is to detect the direction, forward or backward. And the Boat Classification is to identify the boat type, which is the same as bus number. The system is tested with record data from Chao Phraya River Express Boat in Bangkok, Thailand. The result shows that the system can work well, though there is no manual input required.

III. Challenges and Solutions

In this section, we discuss the mentioned challenges and proposed solutions of the automatic transit tracking system using the Boat Arrival Time Prediction System as an example.

A. *Tracking Performance*

In the simplest case, there is only one sharing user on the boat so the accuracy of the system will be based on given information by this user. But if there is more than one user on the boat, the system needs to achieve two goals in this challenge, which are (1) getting the better result, because the system has more information to provide more reliable and accuracy service. And (2) the least battery consumption, since activating the sensors in mobile phone all the time will consume a lot of battery. Therefore, if there are a lot of users on the same boat, it does not necessary that every user has to share their information.

Hence, the solution of this challenge is to find the equilibrium between accuracy and battery consumption. In other words, to find the best technique that can give the best result with the least amount of sharing users. In this paper, we choose scoring technique, which is simple but powerful enough to demonstrate how to solve this challenge.

The scoring technique focuses on finding the best user by scoring and comparing all users on the same boat. To give a proper score, we need to understand how the system works. In our system, the accuracy is based on the Boat-Stop Detection. The detection's purpose is to tell the system that the boat is stopping at a pier and also tell what pier it is. To be able to do that, the detection has to detect whistle sound from boat conductor and Wi-Fi list at the pier correctly. So we can define that the best user is the user that can detect the Wi-Fi list and the whistle sound with the most accuracy.

B. *Tracking Consistency*

In real life situation, passengers can embark or disembark the boat at any piers. So it is possible that the selected sharing user(s) is going to disembark the boat at any piers. Or the new user can also embark the boat at any piers too. The system should be able to handle these two situations efficiently so that it can continue to track the transit smoothly with the best performance.

The first situation, the current sharing user disembarks, can lead to a problem that the transit tracking is lost. To solve this problem, a backup user is introduced. The backup user will wait in a stand by mode, which it will open only necessary sensors to just check that they are still on the boat. And when the sharing user is disembarking, the backup user will become the sharing user to continue the tracking.

The second situation is where a new user embarks into the boat. The system should be able to group the new user together with existing users correctly and evaluates the new user for assigning to be a new sharing user or the backup user.

C. *Tracking Continuity*

The problem of the automatic system comparing with manual input system is that it cannot detect some necessary information instantly, such as Boat Classification technique, which need a sequence of piers that the boat has stopped to classify the type of the boat. This information can be told instantly in the manual input system by the user. The problem becomes more serious when the user can disembark at any piers because the whole information will be gone and the detection has to start again from the beginning. There are three possible scenarios that can happen when the selected sharing user disembarks which are: (1) there is a backup user available on the boat, (2) there is no backup user, but a new user is embarking the boat, and (3) there is no backup user, and a new user does not embark at the current pier but later pier.

The solution for each scenario is different. For the first scenario, the solution is straightforward. The backup user will become the sharing user and continue to complete the detection if it is not completed yet.

For the second scenario, the solution is similar to the first scenario but the system has to make sure that the new user is on the same boat as the previous sharing user. So this means the system has to wait for a while until the same boat can be identified. Since boat environment is not complex, because the boat pier can support only one boat from each different direction at the same time. Hence, the system can easily confirm that the boat is the same by comparing the direction.

For the third scenario, it is more complex than the previous scenarios. When the new user embarks the boat, the system has to check three conditions, which are a list of lost track boats, time, and direction. So every time the user embarks the boat, the system will check if there is any lost track boats. If it has, then check the available information of the boat and compare the time from the pier that the boat lost track until the user embarks the boat with the system's predicted travel time between those piers. And then the direction of the boat should match. Only all three conditions are satisfied, the hand over will be completed.

IV. PERFORMANCE EVALUATION

We evaluate the solutions using Android Phones, Nexus One, to collect the data. Then we simulate the challenges and our proposed solution in our Boat Arrival Time Prediction System, and collect the results.

A. Tracking Performance Solution Evaluation

As we already stated that the best user should be able to detect both Wi-Fi and whistle sound with the highest accuracy. From our experiments, the list of Wi-Fi at any piers is always detected, since the range of outdoor Wi-Fi signal normally can be reached up to 90 meters away. But the device sometimes fails to detect the whistle sound. After we investigated further, we found that the accuracy is varied based on the user's location on the boat. If the location of the user is far from the conductor, the probability that the whistle detection is missed will be high too. So we can conclude that the best user is the user that stays near the boat conductor. However the system cannot tell the user's location on the boat. Therefore, we have to find how loud the whistle is to estimate the distance between user and conductor. Our boat system already has this value since the whistle detection uses Fast Fourier Transform (FFT) so we can look at the whistle Hz and its dB value.

We tested by detecting the whistle sound 5 times with four users in different locations on the boat as shown in Figure 1. The results of the records are summarized in Table I. The results show that if any user stays near conductor, only one sharing user in each boat is enough to give 100% of the accuracy for the Boat-Stop Detection. And the rest of the users in the system can save their mobile phone battery.

B. Tracking Consistency Solution Evaluation

In our simulation for the first situation, we create one sharing user and one backup user in the boat. And let the sharing user disembarks the boat at the next pier. When the system detects that the sharing user is disembarking, it will tell the backup user to change into the sharing user and continue the boat tracking normally without any problem.

Battery consumption is one of the reasons that we introduced the concept of backup user. Therefore, the backup user will perform only single detection, which is Disembark-Detection because the system needs to make sure that the user is still on the boat. The detection operates by using only two sensors, Wi-Fi and accelerometer. Table II compares the sharing user's battery consumption to the backup user's battery consumption, the result shows that the backup user uses battery less than sharing user for 30%, which the battery life can last longer than sharing user for 2 more hours.

The second situation, we simulate one sharing user and one backup user on the same boat. We add a new user embarking at every pier the boat stops. And we also remove the current sharing user at some piers randomly. Using the scoring technique, if the new user's score is better than the current sharing user, the new user will become the sharing user and the

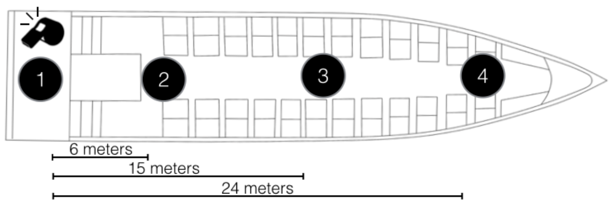

Fig. 1. Boat with recorded location and approximated distance

TABLE I. LOCATION AND WHISTLE DETECTION RESULT

Location	On hand (dBs)	Accuracy	In pocket (dBs)	Accuracy
No. 1	-2	100%	-11	100%
No. 2	-9	100%	-19	100%
No. 3	-11	100%	-29	80%
No. 4	-25	80%	-60	20%

TABLE II. BATTERY DURATION COMPARISON

User type	Battery Consumption in one hour (%)	Battery Duration (in hours)
Sharing user	20	5
Backup user	14	7

sharing user will change to the backup user. But if the new user's score is not better, the user will become the backup user. The system have to store all users' score so when we let the sharing user disembark, the system can change the next best score user to become a new sharing user. From our simulation, the system can track the boat smoothly.

C. Tracking Continuity Solution Evaluation

In the first scenario, where there is the backup user available on the boat, the simulation result shows that the boat information can be handed over from the sharing user to the backup user without any problems and the backup user can continue to track the boat immediately.

In the second scenario, where there is no backup user on the boat but a new user embarks at the pier that the sharing user disembarks. The results show that if there is only one boat stopping at the pier, the system can hand over data correctly and instantly. The time for hand over the data is depended on how long the passengers disembark and embark. Based on our records, it can take 30 seconds up to 5 minutes if the boat and the pier are very crowded. For the situation that there are two boats stopping at the same pier, the system has to wait until the direction of the boat is detected. Since the direction will be known at the next pier, the tracking system can continue without missing any piers.

And in the last scenario, where there is no backup user, no new user embarks at the current pier but will embark in the later pier. From our simulation, we add abilities to the system to be able to store and remove the lost track boats. When the last sharing user disembarks the boat, the system will add the boat to the lost track boat list. And remove it from the list when the boat is matched. The direction of the boat is already detected accurately in our system. So the biggest challenge is time because the time will be used for matching the boat and also can be used for estimate the boat type. To match the boat with the lost track boat, the travel time of the boat from the lost track pier to the current pier should be in an *acceptable range of time*. The *acceptable range of time* is the server's predicted time plus the range of time that the boat normally spends at each pier. For example, if the boat is lost track at Pier 2 and there is a boat appear at Pier 4, the *acceptable range of time* that the boat will match is calculated by summation of travel time from Pier 2 to Pier 4, which is 12 minutes, and the time that the boat spends at Pier 3, which is 30 to 60 seconds. Therefore, the acceptable range of time is between 12.5-13 minutes. To classify the type of the boat, the system uses the difference in the *acceptable range of time* of each boat type. For examples, from Pier 2 to Pier 4, the green flag boat will take only 10 minutes but the orange flag boat will take 12.5-13 minutes. The system can find the type of the boat correctly if it is matched with only one lost track boat. But it is possible that there are more than one lost track boat that are matched because of the difference speed of the different type of boat, the overtake scenario can occur as shown in Figure 2. To solve this problem, the system will let the boat carry all matching boats' information and try to identify the type by the future stop piers.

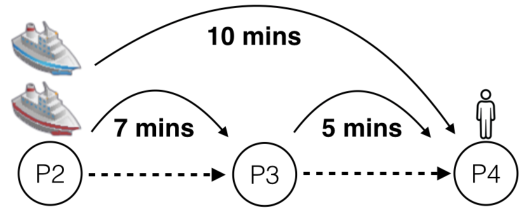

Fig. 2. Overtake situation

V. CONCLUSION

In this paper, we show the challenges of the automatic crowdsourcing real-time transit tracking. We discuss the importance of each challenge and highlight the situations that they will appear. Then, we demonstrate the challenges by simulating them in our boat arrival time prediction system, which is an automatic detection system. We also apply the solutions to those challenges and evaluate them. Though, we show the challenges and the solutions with the boat system, but these concepts can be applied for any transportation, such as bus or train, with some modification based on its characteristic and behavior.

ACKNOWLEDGMENT

This project has been partially supported by Special Task Force for Activating Research (STAR) Funding in Wireless Network and Future Internet Research Group, Chulalongkorn University.

REFERENCES

[1] [Online] Available: http://www.moovitapp.com

[2] A. P. Naveen Nandan, Xing Zhe, "Challenges in Crowdsourcing Real-time Information for Public Transportation," *IEEE 15th International Conference on Mobile Data Management*, 2014.

[3] G. X. Xiang Zhang, Ruozhou Yu, Dejun Yang, Jian Tang, "You Better Be Honest: Discouraging Free-Riding and False-Reporting in Mobile Crowdsourcing," *Globecom 2014 - Wireless Networking Symposium*.

[4] K. N. N. Phuthipong Bovornkeeratiroj, Kultida Rojviboonchai, "Boat Arrival Time Prediction System using Mobile Phone Sensing," 12th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, Hua Hin, Thailand, 2015.

[5] C.-K. T. John Kah-Soon Lau, Tie Luo, "Participatory Cyber Physical System in Public Transport Application," *Fourth IEEE International Conference on Utility and Cloud Computing,* 2011.

[6] J. Z. Aaron Steinfeld, Anthony Tomasic, Daisy Yoo, Rafae Dar Aziz, "Mobile transit information from universal design and crowdsourcing," *Journal of the Transportation Research Board, vol. 2217, pp. 95–102,* 2011.

[7] J. B. Arvind Thiagarajan, Tomas Gerlich, Jakob Eriksson, "Cooperative Transit Tracking using Smart-phones," *SenSys'10,* 2010.

[8] P. Zhou, Y. Zheng, and M. Li, "How long to wait?: predicting bus arrival time with mobile phone based participatory sensing," presented at the Proceedings of the 10th international conference on Mobile systems, applications, and services, Low Wood Bay, Lake District, UK, 2012.