VOL. 11, NO. 12, JUNE 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

Wi-Fi SNIFFER BASED COMMUTERS STATISTICS COLLECTION SYSTEM FOR RELIABLE BUS SCHEDULING SYSTEM

Kim Chuan Lim, Jian Ea Hoo and Yin Hui Lee
Faculty of Electronics and Computer Engineering, Universiti Teknikal Malaysia Melaka, Malaysia
E-Mail: kimchuan@utem.edu.my

ABSTRACT

Traffic congestion issues have always been a concern for the fast growing metropolitans in which more than 90 percent of trips are made entirely by private means of transportation i.e. by car and motorcycle. As the country is actively engaged in infrastructure development especially in the transportation network to facilitate the movements of people and goods, a high demand for better public transportation is needed to reduce the issue of road congestion (percentage of GDI lost due to man hour lost in the traffic). Therefore, a cost effective Wi-Fi sniffing based bus commuters' statistic collection system is designed and developed to study the feasibility of predicting the necessity of scheduling additional bus services when the detected number of Wi-Fi enabled devices exceeded the bus capacity. The developed system is subsequently deployed to the busiest university bus stop and the obtained result shows that variation of sniffed MAC address exhibit parallelism to the actual number of commuters waiting at the bus station as observed in the captured bus station video images. Result also shows that the MAC address based counting system can help to alert the bus management for better scheduling when the commuter at the particular bus stop is traveling to the same destination.

Keywords: Wi-Fi sniffing, raspberry pi sniffer, unique MAC addresses based commuter statistic collection system.

INTRODUCTION

Traffic congestion happens every day in megacity and this will certainly contribute to a loss of productivity. Alternatives measures such as MRT project, building new highways and widen the roads have been implemented by the government. However, the proposed solution accounts for large funding. The possible solution is to make use of the available resources such as upgrading the existing public transportation services to overcome the traffic congestion on roads. Bus services have always been an issue for both the university management and the students commuting between the university and the hostel. The location of the students' hostel is hardly reached by walking distance; largely dependent on the bus and other mode of transportation. However, the ineffective bus system has always led to various problems encountered by the students. For instance, the bus capacity can only afford 40 students while at some particular hours, the number of students exceeded the carrying capacity. Time is wasted when there is no additional bus provided and bus that is overcrowded just make the commuting experience even worst.

A quick observation at the busiest bus station in the main campus of the university (Universiti Teknikal Malaysia Melaka) revealed that most of the students are browsing the Internet while waiting for the scheduled bus to be arrived. If a number of them are using the free Wi-Fi internet access provided by the University while waiting for the bus to arrive, a Wi-Fi sniffing based bus commuters destination prediction system, counting the number of unique Wi-Fi MAC address embedded throughout the communication between the smartphone and the access point, could help to predict the need of scheduling additional bus whenever is needed.

This paper is organized as follows. In section 2, the review on different technologies used by existing bus commuters' statistic collection system is discussed. In section 3, methods used to design and deploy the Wi-Fi sniffer based commuters' statistic collection system is presented. The experimental results of deploying the proposed system are presented in section 4. Finally our work of this paper is summarized in the last section.

Review on technologies used by existing bus commuters statistic collection system

The Wi-Fi MAC address uniquely possessed by the smartphone and the signal strength measured by the Wi-Fi adapter are the two common values that could be sniffed when the smartphone started any Wi-Fi communication. Counting the number of unique MAC addresses could help to identify the number of surrounding smartphones while the measured signal strength could roughly provide the proximity of the smartphone with respect to the sniffer. Making use of this information, novel applications such as building occupancy and crowd data collection [1,2,3,4], pedestrian activity monitoring [5] service queue length monitoring [6] and queuing time prediction [7] have been made possible, thanks to the widely available of low power Linux Wi-Fi stack enabled single board computers in the market.

Easing the traffic congestion required the citizen in the city to give up the traveling with personal vehicle and use the available public transport. Projects on data-driven real time monitoring and incident management of public transport in the city [8, 9] have become the top priority of the city manager to provide a reliable and smart public transport service to gain the confidence of the commuters.

It is undeniable that when dealing with public transportation, time and patience are in paramount consideration. To develop a smart bus system, different approaches likes scanning QR code using the smartphone with the Quick Response code reader, passengers can predict the bus arrival time, view the buses' locations and bus routes on a map [10]. Conventional bus commuter statistic collecting system would involve the use of electronic ticketing system [11] and video image processing [12] to obtain the required data for subsequent bus schedule planning. Unfortunately, most of these systems are installed in the bus, only able to improve the bus scheduling in the future, but not provide real time bus stop commuter statistic which can help prediction the necessity of dispatching additional bus services when needed.

DESIGN AND DEVELOPMENT OF THE BUS COMMUTERS STATISTIC COLLECTION SYSTEM (Wi-Fi SNIFFER WITH RASPBERRY PI)

The Raspberry Pi 2 Model B is one of the most affordable embedded boards which consists of Quad-CPU and 1GB of RAM. Using the Rasperry Pi 2, together with USB Wi-Fi dongle, Camera, and power bank, a Wi-Fi sniffer based bus commuters statistic collection system is designed and developed as shown in Figure-1. The developed system is subsequently deployed to the busiest bus stop in the university as shown in Figure-2.

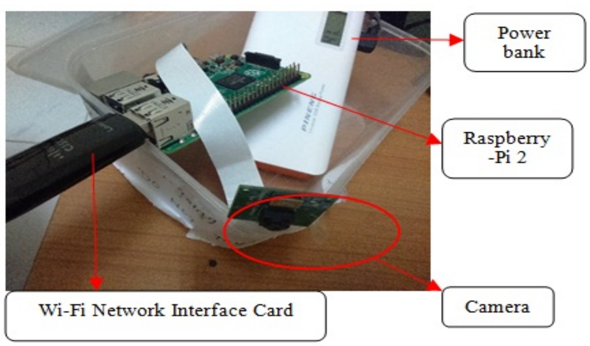

Figure-1. Prototype of the designed and developed bus commuters' statistic collection system.

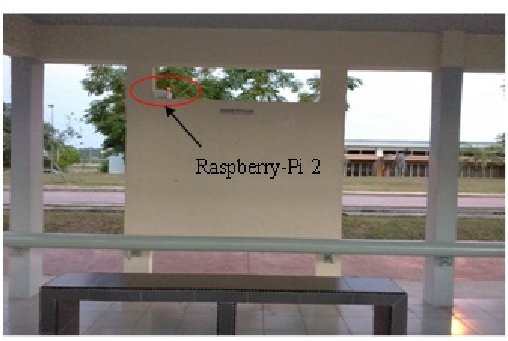

Figure-2. Bus commuters statistic collection system (Raspberry Pi based Wi-Fi sniffer) deployed at the bus stop.

A camera module is connected to the Raspberry-Pi 2 and configured accordingly with the built in tools (raspi-config). A python script is created to allow a camera image to be captured every two seconds and stored into the SD Card (the size of an image is around 250Kbyte. This will require approximate 2GByte of storage for 8 hours of recording). Each of the images is stored according to the time it is being captured.

Python-Scapy [13] is identified as the sniffing software library tool for this project due to its support for direct capture and process of various types of Wi-Fi frames and their sub-field through the built-in RadioTap header mechanism.

Pre-filtering of the sniffed frame, logging only the probe request frames directly contributing to the bus commuter analysis, is required to prevent from logging all the information (huge storage needed). To avoid sniffing too many repeated data, time based pre-filtering has been applied. For example, the same MAC address which appears more than five seconds from the preceding discovery will be recorded again, otherwise, it will not be stored into the log files. Post processing of the data is then carried out.

Bunga Raya, Emerald Park and Sri Utama are the three off campus hostels, around 12KM from the UTeM main campus, where lots of UTeM students are staying. Students staying in Bunga Raya and Emerald Park will be traveling to the main campus with the same bus service (Bunga Raya hostel is only 1KM away from the Emerald Park). Most of the times a second bus will be manually scheduled to send the student back to the two hostels during the peak hour.

Observed unique MAC address study at the bus stop will be performed 15 minutes before the expected arrival of the bus (based on the given UTeM Bus Timetable) and 15 minutes after the scheduled bus (in case the bus is late). The observed unique MAC address will be processed and grouped into 5 minutes interval to allow effective presentation of the accumulated unique MAC address (the connecting and disconnecting of the smartphone Wi-Fi connection of the passenger to the UTeM access point will not impact the result of the study).

The sniffing results obtained from probe request frames are processed using Python script to remove repeated MAC address for the interval of five minutes and presented in 5 minutes bar chart interval to ease the bus prediction analysis

EXPERIMENTAL RESULTS

The amount of the MAC address collected across the experiment day are manually cross validated with the captured camera image at different time intervals according to the scheduled bus timetable. The number of sniffed MAC addresses and the observed number of passengers waiting at the bus station boarding to Emerald Park and Sri Utama are tabulated as shown in Table-1 and plotted with a bar chart (15 minutes before and 10 minutes after the expected bus schedule) as shown in Figure-3. The obtained results are discussed as follow.

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

Table-1. Difference between number of MAC address sniffed correspond to the actual number of passengers observed from captured image.

Observed bus arrival time (Emerald Park)	Observed bus arrival time (Sri Utama)	Number of MAC address obtained by sniffing probe request frames	Observed number of passengers to Emerald Park	Observed number of passengers to Sri Utama
11:34am	11:30am	30	25	3
12:30pm	12:31pm	27	23	4
1:31pm	1:32pm	22	9	6
2:31pm	2:34pm	32	15	14
4:31pm	4:34pm	57	39	13
No bus observed	5:27pm	57	47	17
Bus #1 6:30pm Bus #2 6:34pm	6:33pm	51	60	2

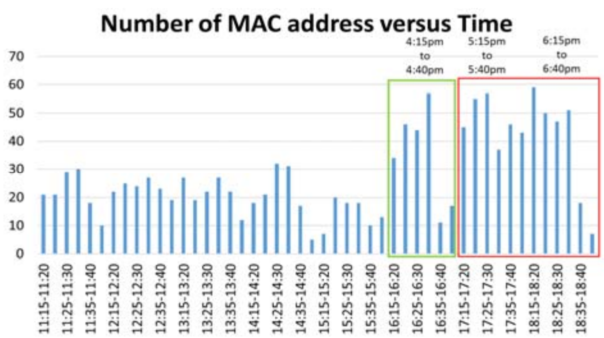

Figure-3. Number of unique sniffed MAC address vs time (15 minutes before and 10 minutes after the expected bus schedule) of the selected experiment day.

Scheduled bus: 2:30pm (sniffed 32 unique MAC addresses, observed 29 commuters)

Figure-4 shows the captured image revealing the actual number of passengers waiting at the bus station, destined for Emerald Park, according to the scheduled bus timetable. Passengers waiting at the bus station, whereby their destination is Sri Utama is shown in Figure-5.

Figure-4. Bus going to Emerald Park arrived at 2:31pm. observed number of commuters from the recorded image sequences is around 15 persons.

Figure-5. Bus going to Sri Utama arrived at 2:34pm. Observed number of commuters from the 1 minute recorded image sequence is around 14 persons.

Scheduled bus: 4:30pm (sniffed 57 unique MAC addresses, observed 52 commuters)

The sniffed MAC address peak is shown at 4:30pm (see the Figure-3 bar chart). The number of MAC address sniffed reached 57. The results are cross validated with the image capture where the observed actual number of passenger is estimated around 52. Among these 52 passengers, 39 of them are observed to be traveling to Emerald Park (refer Figure-6). The remaining ones are observed to be traveling to Sri Utama (Figure-7). Even though the number of sniffed MAC address and the actual number of passengers observed from the captured image will most probably exceed the maximum bus capacity, no additional bus shall be scheduled.

Figure-6. An estimate number of 39 commuters observed from the captured image at 4:31pm are traveling to Emerald Park.

Figure-7. 17 observed number of commuters traveling to Sri Utama at 4:34pm.

Scheduled buses: 5:30pm (bus service disruption, bus service restored at 6:30pm)

Based on the captured camera image sequence, the number of students waiting at the bus station at 5:30pm are estimated to reach 65 (manually counted, refer to Figure-8). Figure-9 shows that the students waiting at the bus station at 6:30pm are the same batch of people waiting at 5:30pm. This indicated that there is no bus service provided for the students at 5:30pm even though there should be bus service according to the scheduled bus timetable. The bar chart (see Figure-3) is showing total number of sniffed unique MAC addresses exceeded 40 from 5:30pm onward.

Figure-8. Observed number of commuters waiting at the bus station at 5:30pm is estimated at approximate of 65.

Figure-9. Observed number of commuters waiting at the bus station at 6:30pm is estimated at approximate of 62 whereby most of the passengers have been waiting for an hour (from 5:30pm to 6:30pm).

CONCLUSIONS

From the obtained result, the variation of the sniffed MAC address exhibit parallelism to the actual number of commuters waiting at the bus station (total number of sniffed unique MAC addresses variate with respect to the commuter observed in the captured camera image sequence). Additional historical MAC address tracing should be able to help eliminate the false alarm to schedule additional buses when the total sniffed unique MAC addresses exceeded 40 (bus scheduled 4:30pm). The obtained results also show that the designed and implemented system can be used to alert the bus scheduling management when lots of commuters are waiting for the bus to be scheduled (bus scheduled 5:30pm).

Wi-Fi sniffing based commuter destination predicting system installed at the bus station that is within the free Wi-Fi coverage, where commuters are enjoying the free Internet services while waiting for the arrival of bus, could definitely assist in generating a commuters' statistic to help predict, in real time, the need of scheduling additional bus which could reduce the waiting time and improve the commuting experience.

ACKNOWLEDGEMENTS

This work is supported by Knowledge Transfer Programme (I-ECO/32 (UTEM-14)).

REFERENCES

[1] Abedi N, Bhaskar A & Chung E. 2013. Bluetooth and Wi-Fi MAC Address Based Crowd Data Collection and Monitoring: Benefits, Challenges and Enhancement. Australasian Transport Research Forum 2013 Proceedings.

[2] Qin WJ, Zhang JD, Bo L, Sun LM. 2013, Discovering Human Presence Activities with Smartphones Using Nonintrusive Wi-Fi Sniffer Sensors: The Big Data Prospective. International Journal of Distributed Sensor Networks.

[3] Radu L, Kriara L & Marina MK. 2013, Pazl: A Mobile Crowdsensing based Indoor WiFi Monitoring System. Network and Service Management. 9th International Conference on IEEE. Page 75-83.

[4] Ruiz AJ, Blunck H, Prentow TS, Stisen A & Kjaergaard MB. 2014, Analysis Methods for Extracting Knowledge from Large-Scale Wi-Fi Monitoring to Inform Building Facility Planning. 2014. IEEE International Conference on Pervasive Computing and Communications. Page 130-138.

[5] Zhuliang X *et al*, 2013. Pedestrian Monitoring System using Wi-Fi Technology and RSSI Based Localization. International Journal of Wireless & Mobile Networks (IJWMN). Vol.5. No. 4.

[6] Wang Y, Yang J, Chen YY, Liu HB, Gruteser M & Martin RP. 2014, Tracking Human Queues Using Single-Point Signal Monitoring. Proceedings of the 12th annual international conference on Mobile systems, applications and services. Pages 42-54.

[7] Stephen Lawson. 2014. At Austin Airport. Wi-Fi Predicts How Long the Security Line will be. [Online]: http://www.itworld.com/article/2837781/at-austin-airport-wi-fi-predicts-how-long-the-security-line-will-be.html

[8] Land Transport Authority of Singapore. 2014. LTA, SMRT, Starhub and IBM collaborate to improve transport with data for Singapore commuters. [Online]: http://www.lta.gov.sg/apps/news/page.aspx?c=2&id=407a5053-0345-40f5-8d64-51fb31bfb2a0

[9] IMB Press Room. 2012. City of Zhenjiang, China and IBM Collaborate to Build a Smarter City. [Online]:https://www03.ibm.com/press/us/en/pressrelease/36883.wss#release

[10] S. Eken & A. Sayar. 2014. A Smart Bus Tracking System Based on Location-Aware Services and QR Codes. Innovations in Intelligent Systems and Applications Proceedings, 2014 IEEE International Symposium. 299-303.

[11] Kimpel, Thomas, James Strathman, David Griffin, Steve Callas, and Richard Gerhart. 2003. Automatic passenger counter evaluation: Implications for national transit database reporting. Transportation Research Record: Journal of the Transportation Research Board 1835. 93-100.

[12] Chen, Chao-Ho, Yin-Chan Chang, Tsong-Yi Chen, and Da-Jinn Wang. 2008. People counting system for getting in/out of a bus based on video processing. In Intelligent Systems Design and Applications. ISDA'08. Eighth International Conference on. IEEE. vol. 3. pp. 565-569.

[13] Bansal, S., and N. Bansal. 2015. Scapy–A Python Tool For Security Testing. J Comput Sci Syst Biol 8. 140-159.