How does coffee shop get crowded?: Using WiFi footprints to deliver insights into the success of promotion

Pichaya Prasertsung
School of Information Computer and Communication Technology,
SIIT, Thammasat University,
Pathumthani, Thailand
m5822040068@studentmail.siit.tu.ac.th

Teerayut Horanont
School of Information Computer and Communication Technology,
SIIT, Thammasat University,
Pathumthani, Thailand
teerayut@siit.tu.ac.th

ACM ISBN 978-1-4503-5190-4/17/09...$15.00
https://doi.org/10.1145/3123024.3124418

Abstract
Real time people density estimation is one of the biggest challenges of today research. This information can be applied to various urban applications such as advertising,traffic planning, and resource management. The recent researches have demonstrated on crowded estimation using various cutting-edge technologies to solve this problem. Since WiFi access points already exist in the major buildings and shops. They become a great tool to estimate the density of people. This research explores the opportunities to use existing access points in order to estimate density of people in the real world environment. WiFi probe request monitoring technique is used to identify the number of the customer's visit to a coffee shop. The result shows that during weekdays, the number of customers in promotion period is 30.43% greater than non-promotion period.

Author Keywords
Retail analytics; WiFi footprint; Mobile devices; Geolocation; Knowledge discovery;

ACM Classification Keywords
H3.3 Information Search and Retrieval: information filtering

Introduction

Knowing the number of people in a specific area can be applied to various applications. For events, organizers can prepare an evacuation plan in a case of emergency. Shop owners can develop marketing plans based on the number of customers visiting their stores. For reasons above, various technologies are applied for counting people as explained in next section.

Nowadays, people tend to carry their smartphone everywhere they go. For this reason, we can estimate the number of people based on the amount of detected smartphones especially in the dense urban areas. There is a study on estimating the number of people using only RSSI at the node from the existing WiFi access point. By testing in experimental environment, the accuracy rate for estimating the degree of congestion is 0.946 [1]. With this acceptable result, this research focuses on exploring the efficiency of monitoring public WiFi signal to detect the number of smartphones at the café. The weekday discount campaign is use to demonstrate the performance of this work.

The rest of this paper is organized as follows. Background and related technologies for people density estimation are explained in section 2. Data collection and data analysis are described in section 3. Results and discussions are explained in the section 4. The last section, section 5, will conclude and discuss about limitation, and future work for this researches.

Background & Related works

Various technologies are selected to estimate the people density. There are 2 common technologies; computer vision based technique and non-computer vision based technique. For computer vision technique, camera is used to capture the environment. Then features are extracted from videos or pictures to analyze in order to estimate the number of people. The drawback of this approach is that the accuracy is drop in dim environment [2].

For non-computer vision technique, various sensors such as GPS, WiFi, and Bluetooth can be applied. GPS is a good choice to estimate people density in outdoor environment. Horanont et al [3] states that deep analyze mobile phone GPS data can help urban activities and their evolution through space and time using a case study of Odaiba area in Japan. Additionally, Ricciato et al [4] presents that mobile phone network-based data can be used to estimating population density distribution. However, GPS is not efficient to estimate people density in the indoor environment because the signals from satellites is attenuated by obstacles such as roofs and walls.

For indoor environment, WiFi is a suitable option because it requires no additional hardware installation. Using WiFi to track people can be classified to 3 techniques as follows

1. WiFi RSSI measurement
2. Channel state information (CSI)
3. WiFi probe requests and responses

For WiFi - RSSI technique can be used to count the number of people by measuring signal transmit and receive between antennas. Machine learning can be applied to classify the result of signal. Depatla et al [5] suggested that number of people walking in an area could be counted by measurements WiFi RSSI between a pair of stationary transmitter and receiver antennas. Yoshida and Taniguchi [5], confirmed the accuracy of

Laptop information for this experiment

Features	Information
Model	MacBook Pro (13-inch, Mid 2012)
Processor	2.5 GHz Intel Core i5
Memory	4 GB 1600 MHz DDR3
WiFi Card type	AirPort Extreme (0x14E4, 0xF5)
WiFi Supported PHY Modes:	802.11 a/b/g/n
WiFi Supported Channels:	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 36, 40, 44, 48, 52, 56, 60, 64, 100, 104, 108, 112, 116, 120, 124, 128, 132, 136, 140, 149, 153, 157, 161, 165

Table 1: Laptop information

estimating the presence and absence of people is up to 0.982 using existing WiFi access point in indoor environment. Another technique is to use CSI. Zeng et al [6] proposed an approach that leverages CSI from WiFi network to analyze customer behavior. The last technique is to monitor WiFi probe request and response between devices and access points. Schauer et al [7] report that WiFi tracking provides a good approximation to crowd densities and pedestrian flows combining with additional related data such as opening times of the security gate.

This research, we demonstrate the use of WiFi probe request technique with our developed filtering methods to accurately count the customers during weekday where the campaign period begin at the University's coffee shop.

Methodology

In this implementation, we process 3 steps to confirm our results as follows

1. System setup

For this work, we install software on our Macbook and count the number of mobile devices by monitoring unique MAC address of proximity cellphones. Refer to the finding of Taylor et al, 70% of smartphone owners connect to internet via WiFi, thus, the number of people could be calculated as

$$N_{people} = \frac{N_{devices}}{0.7}$$

This software is running on MacBook pro; Sierra 10.12.3. Detailed information about laptop is described in table (1)

2. Information Collection

From exploration, we monitor WiFi by running a program to capture probe requests between mobile devices and wifi access points. On laptop, WiFi adapter is turned to monitor mode to sniff the probe requests. Mac Address, RSSI, and capturing time are collected. At first, we start monitoring a day that this cafe offers a promotion. Customers will get 50% discount when they buy drinks from 14:00 to 17:00. The interval for capturing is 60 seconds. Figure 1 shows the number of detected devices over time.

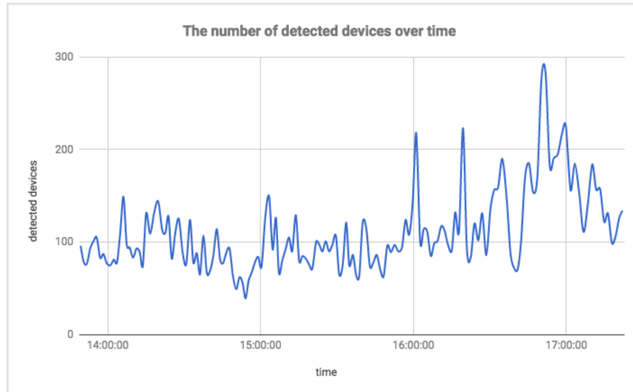

Figure 1: Raw data of the number of devices that the system can detect over time.

Apart from monitoring probe requests, we manually count the number of customers in cafe. Then we classify the number of people into 3 classes as follows.

Figure 2: The scenario of this cafe during the campaign period.

Figure 3: The scenario of this cafe during weekday. There is no promotion on this day.

Class	Definition (1)	Definition (2)
Less	Less than 40 customers	Less than 28 devices found. (70% of 40 is 28)
Medium	40 - 60 customers	28 - 42 devices found.
High	More than 60 customers	More than 42 devices found.

Table 2 A class to estimate people density in store. Since 70% of devices enables WiFi [8], the number of detected devices is multiplied with 0.7

From observation, the number of customers is considered as *High* during 14:05 -14:20, 15.30–16.00, and 16:40 - 17:00. After 18.00, the number of people is decreasing dramatically which consider as *Less*. On that day, it was rain from 14.30 - 15.00. Apart from the time referred above, the number of customers is considered as *Medium*. Figure 2 displayed the scenario of the café during the campaign period. Figure 3 displayed the scenario of the average weekday customers.

3. Data Processing

In this step, we ignore the data which RSSI greater than -70 dBm because we want to focus only nearby devices, more straightforward, only the customers in the coffee shop. From 13:30:00 - 17:30:00, the number of detected devices before filtering out RSSI is 8,934 devices. The number of devices, which RSSI greater than -70 dBm, is 1,011 devices.

Due to that fact that customers spend more than a minute in this café, the devices that are detected only one time by the system will be ignore as shown in table 3.

Criteria	The number of devices	Detection Percentage
All detected devices	8934	100
Detected devices with RSSI greater than -70 dBm	1011	11.32
Detected devices with RSSI greater than -70 dBm and the number of occurances is greater than one	170	1.90

Table 3: Number of detected devices over each criterion.

Discussion & Results

From observation, the number of customers in the café before the promotion start is *Medium*. There are less than 5 customers stand at the counter to order drinks. Empty seats are available for new customers. However, 5 minutes after the promotion starts, many customers come. They wait to buy drinks at a counter during 14.05 to 14.20. After 14.30, the number of customers becomes *Medium* again. Around 16.40, many customers visit the café again. All seats in café are already occupied. There are more than 10 people at the counter waiting for orders. Figure 4 displays the average number of people on a promotion day and non-promotion day over time period. From the graph, number of people is significantly increasing before 17.00, which is the time before the promotion end.

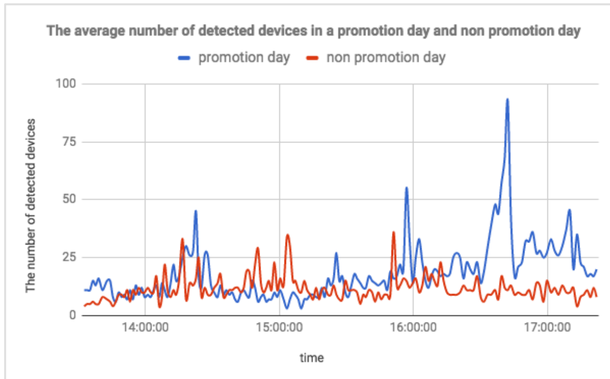

Figure 4: The total number of detected devices on a promotion day and non-promotion day

Based on information above, we separate data into time slot with interval of 15 minutes. Devices with only one time detection in each period will be ignored. Figure 5 displays the number of detected devices with the interval of 15 minutes on a promotion day and non-promotion day during the campaign.

From observation, the number of customers increases at 14:05 due to the promotion time. This relates to the increasing of detected devices as shown in line graph in figure 5. The number of detected devices starts increasing from 18 devices to 26 devices before the promotion start. Then, the number of devices significantly increase again one hour before the promotion end.

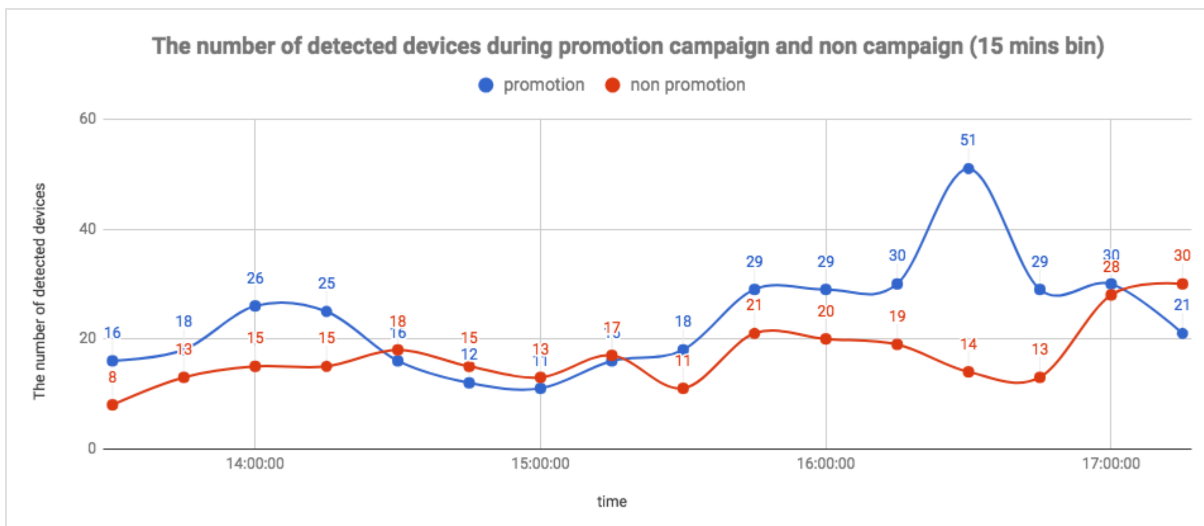

Figure 5 The number of detected devices on a promotion day and non-promotion day using 15 minutes bin when ignore the device that can be detected only one times.

From the graph, the average detected devices, which found on a promotion day and non-promotion day, are 23 devices and 16 devices, respectively. The average number of detected devices is 30.43% increase on promotion day. In the first period of the campaign from 14.00 – 14.30, the number of detected devices is 42.31% increase compared with the same period without promotion. Additionally, an hour before the promotion ends, from 16.00 – 17.00, the number of detected devices is 52.52% increase compared with the same period on non-promotion day.

Conclusion and Suggestion

Monitoring WiFi probe requests can decently estimate in a coffee shop. The result shows that the number of customers tends to increase on the average of 30.43% on a promotion day. This can be used to explore how a promotion can drive customers to your stores. Filtering devices with outside the store can be done by limit the number of known access points. For example, only the number of devices, which connected to the store's WiFi access point, is considered. This kind of data can estimate the density of people over time without installing additional hardware. The results can be used to develop a marketing plan to increase profit of the stores. Explore density in different environment and integrate this technique with other solution can be done for further research. Develop algorithm to remove noise or outlier can be done in order to improving accuracy of density estimation.

Acknowledgement

Part of this research is supported by Thammasat Young Researcher Fund 2016.

References

1. Yoshida Takuya, and Yoshiaki Taniguchi. Estimating the number of people using existing wifi access point in indoor environment. In *Proceedings of the 6th European Conference of Computer Science (ECCS'15)* (Rome 2015), 46-53.

2. Julio Cezar Silveira Jacques Junior, Soraia Raupp Musse, and Claudio Rosito Jung. Crowd Analysis Using Computer Vision Techniques. *IEEE Signal Processing Magazine*, 27, 5(September 2010), 66- 77.

3. Teerayut Horanont, Santi Phithakkitnukoon, Ryosuke Shibasaki. Sensing Urban Density Using Mobile Phone GPS Locations: A Case Study of Odaiba Area, Japan. *Nature of Computation and Communication. ICTCC* , 144 (January 2015), 146-155.

4. Ricciato Fabio, Peter Widhalm, Massimo Craglia, and Francesco Pantisano. *Estimating population density distribution from network-based mobile phone data*. technical report 978-92-79-50193-7, Publications Office of the European Union, 2015.

5. Saandeep Depatla, Arjun Muralidharan, and Yasamin Mostof. Occupancy Estimation Using Only WiFi Power Measurements. *IEEE Journal on Selected Areas in Communications*, 33, 7 (July 2015), 1381 - 1393. http://ieeexplore.ieee.org/document/7102673/

6. Yunze Zeng, Parth H. Pathak, and Prasant Mohapatra. Analyzing shopper's behavior through wifi signals. In *the 2nd workshop on Workshop on Physical Analytics* (Florence 2015), 13-18. http://dl.acm.org/citation.cfm?id=2753508

7. Lorenz Schauer, Martin Werner, and Philipp Marcus. Estimating crowd densities and pedestrian flows using WiFi and bluetooth. In *the 11th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services* (London 2014), 171-177. http://dl.acm.org/citation.cfm?id=2693006

8. Andy Young, and Andy Noronha Stuart Taylor. (2012, May) Cisco. What Do Consumers Want from WiFi?. Retrieved May 1, 2017 from http://www.cisco.com/c/dam/en_us/about/ac79/docs/sp/SP_WiFi_Consumers.pdf