MiamiMapper: Crowd Analysis using Active and Passive Indoor Localization through Wi-Fi Probe Monitoring

Nicholas Jarvis, John Hata, Nicholas Wayne, Vaskar Raychoudhury, Md Osman Gani
Miami University
Oxford, Ohio, USA
{jarvisna,hatajm,waynens,raychov,ganim}@miamioh.edu

ABSTRACT

Crowd analysis and monitoring in large congregations is an important problem related to public safety and planning. Researchers have long been using image and video for effective crowd analysis. However, with the advent of more sophisticated technologies, crowd monitoring has been attempted using GPS, RFID as well as Bluetooth based systems. Indoor crowd monitoring in public buildings, such as, airport, museum, theaters, etc. has not been widely studied. In this paper, we propose a non-invasive Wi-Fi based approach for indoor crowd analysis which works by passively monitoring the probe packets generated by Smartphones. Since, smartphones nowadays can be unequivocally associated with a human user, counting unique MAC addresses of smartphones can generate a fairly close estimate of number of users in the indoor space and their movement patterns. Although indoor localization using wireless technology has been extensively studied, using probe packets for crowd monitoring has rarely been investigated. In this paper, we have developed the *MiamiMapper* system, from COTS hardware and software, for crowd analysis in indoor environment using passive monitoring of probe packets emitted by wireless devices. Our system can localize users and thereby detect a crowd with an average accuracy of 7.28 meters which is 75.33 % more accurate than GPS based systems for indoor environment.

KEYWORDS

Crowd Analysis, Indoor Localization, Packet Sniffing, MAC, Wi-Fi

ACM Reference Format:
Nicholas Jarvis, John Hata, Nicholas Wayne, Vaskar Raychoudhury, Md Osman Gani. 2018. *MiamiMapper*: Crowd Analysis using Active and Passive Indoor Localization through Wi-Fi Probe Monitoring. In *Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY*. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

With the stupendous growth in smartphone usage various novel research directions have opened up, such as, providing location-aware services, human identification and tracking etc. for both indoor and outdoor environment. Location Based Services (LBS) can be benefited by monitoring and studying variation of crowd density with time as well as their movement patterns, known as crowd analysis. Crowd analysis in indoor environment (such as airport, shopping malls, stadiums etc.) can have huge economic impact as it is possible to build the knowledge of the shops / locations visited by users and the time spent there. Also, the favorite location of users can be identified, and proper infrastructure set up can be arranged to cope with the popular demand. Additionally, several mathematical models can be developed based on long term data for classifying crowd density levels and to predict them. Building crowd forecasting models can be useful for event detection and decision support for proposing mitigative actions during sudden occurrence of high risk incidents like fire or violence.

Although crowd analysis and monitoring is not a novel problem, the existing solutions applied mostly image and video based analysis techniques, grossly known as vision-based methods [29]. These methods are unsuitable for real-time processing required for immediate decision making. Moreover, existing localization systems/techniques [28][15][27] are aimed mostly at individual level and not for groups of people. Although some works have been done for tracking and monitoring large crowd gatherings for Hajj [2][22][21][23] and Kumbh pilgrims [17] they are mostly using various proprietary GPS-enabled RFID readers for outdoor localization or a combination of GPS and Wi-Fi techniques for combined indoor-outdoor localization. Moreover, localization using proprietary products [18][16] has serious shortcomings in terms of compatibility and integration with their peers.

In this paper, we plan to achieve crowd monitoring in indoor environment through passive sniffing of Wi-Fi probe requests generated by the client devices. Our aim is to localize multiple users (crowd) and to track their spatio-temporal pattern of movement over a longer term. Although indoor localization is a well-researched problem [19][10][33] and Wi-Fi based indoor positioning has also been deeply investigated [34], *localizing users for crowd analysis through passively monitoring the Wi-Fi probe requests* appears to be a fairly novel problem worthy of deeper investigation. Wi-Fi probe requests are special network frames emitted by wireless client devices to discover 802.11 network Access Points (APs) in their vicinity. Probe requests are transmitted (although infrequently) even when the device is already connected to an AP and even if the screen is turned off (not in active mode). Since users nowadays carry their smartphones almost all the time and everywhere, we can fairly logically associate a user's location with that of their smartphone. As the probe requests carry the unique MAC addresses of users (MAC randomization is so far used in very limited scope), this enables us to track a user in indoor environment, if we can track the location of their smartphone.

Woodstock '18, June 03–05, 2018, Woodstock, NY

ACM ISBN 978-1-4503-9999-9/18/06...$15.00
https://doi.org/10.1145/1122445.1122456

User localization in indoor environment can take place either through an active localization approach or a passive one. By *active localization*, we refer to the location fingerprinting technique [12] which works by logging the RSSI values of multiple nearby wireless Access Points (APs) observed by the user's smartphone. Multiple phones can be used to generate a more robust fingerprint as the wireless adapters vary in their capabilities to receive signals from APs. Also, RSSI values vary over the days of the week as the crowd density changes inside the building. The active localization approach uses the previously generated location fingerprints to estimate the location of a new user w.r.to the RSSI values observed by his/her smartphone. This approach is not novel and has been widely used and researched. *Passive localization*, on the other hand, deploys a set of listeners which passively listens to the Wi-Fi probe requests generated by the smartphones around them. By parsing the probe requests they estimate the number of users and their locations.

In this paper, we have developed and deployed a system which uses both active and passive user localization in indoor environment and carries out detailed analysis of smartphone user crowds. Having both the approaches of localization using the same system enables us to compare their accuracy and efficiency even though our main purpose is to establish the usefulness of the passive localization approach for monitoring and modeling large indoor crowd dynamics. Our system is developed with COTS hardware and software incurring minimal cost with maximum benefits to the administrators. Although the approach of listening and capturing probe requests been criticized [14] for breaching privacy of individuals, with proper monitoring and MAC address anonymization, user identities can be protected from misuse.

Our proposed system is called *MiamiMapper* which works by capturing Wi-Fi probe requests captured through a series of Static Sensing Units (SSUs), each comprising of a Raspberry Pi module and a powerful Wi-Fi adapter. The SSUs are deployed in strategically important areas of a three-story-building covering all entry and exit points and they passively scan the probe packets transmitted by nearby Wi-Fi enabled devices and extract their unique MAC ids in order to localize the device owners using a pre-calculated location fingerprint. Our proposed approach can be useful for smart building management in terms of energy saving [1], emergency evacuation [31], infrastructure planning, etc. It can also be useful for planning public transport services [8][26] based on long-term crowd monitoring in and around public buildings, such as airports, shopping malls and museums. Overall, our proposed system has significant economic and logistic implications for the business sectors. In summary, the main contributions of this paper are as follows.

- We propose *MiamiMapper*, a prototype system for crowd monitoring and analysis in indoor environment which works by passive capturing of Wi-Fi probe requests using COTS hardware and software systems.
- Our system localizes multiple users with high accuracies through both an active and a passive localization approach and compares their performances. Locations of user crowds are then plotted in a real time basis on a indoor map of a building to generate a heatmap.
- Our extensive experiments with a prototype testbed shows that our proposed system is efficient, scalable, and 75.33% more accurate than GPS based indoor localization techniques.

2 RELATED WORK

Crowd monitoring research, thus far have mainly focused on image or video sequences [29]. However, the drawbacks of using such systems for large-scale real-time situational awareness is their slow and inefficient processing [9]. With the advent of better quality sensory equipment the research in crowd monitoring and analysis have moved towards non-vision approaches, like smartphone, GPS, RFID and Bluetooth based approaches [5][30][32].

There are plenty of research in indoor localization and tracking which always encouraged researchers to achieve higher accuracy with lower cost as well as response time [20]. There are various existing technologies for user localization focusing on single positioning / wireless technology, such as RFID [24], Wi-Fi enabled devices [34], Bluetooth (BT) / Bluetooth Low Energy (BLE) tags [7], and Smartphone's GPS/ GPRS [6] and they can track users in indoor as well as outdoor environment. RFID-based systems suffer from high cost and low performance issues because of human body interference. Pure Wi-Fi based techniques mainly use location fingerprinting which is environment-dependent and unstable as the use Received Signal Strength Indicator (RSSI) values from wireless mobile devices or access points change frequently due to various reasons, such as, signal multipath / interference caused by the interior setup and blocking signals. Multiple high-quality survey papers have been authored on wireless indoor localization systems [19][10][33][12] with in-depth description and analysis of the existing systems.

Localization techniques using the cellular networks achieve good accuracy only in presence of multiple base stations in the vicinity. Similarly, using multiple location technologies together can improve localization accuracy, like combined use of GPS-Wi-Fi-Cellular techniques or GNSS along with Bluetooth, NFC, inertial sensors [25].

Tracking and localizing a large crowd have been experimented w.r.to festivals like Hajj and Kumbh where thousands and even millions of people congregate for religious purposes. A GPS coordinate tagged RFID based localization scheme has been proposed by Ali, et al. [2] for Hajj pilgrims. Similar other systems using RFID-capable smartphones capturing user locations and identities and uploading to remote servers have been proposed [22][21][23]. A combined indoor outdoor seamless localization technique using mobile stations for smartphone probe capturing was proposed in [17]. Their system uses a hybrid of GPS Wi-Fi and BLE probe based localization technique for identifying, tracking and managing people in large crowds congregated on religious purposes during the Kumbh fair in India. However, their paper does not focus on crowd analysis or monitoring but studies mainly seamless localization across indoor and outdoor spaces.

Localization using captured Wi-Fi probes have been examined by Barbera, et al. [4] with a bid to identify the crowd dynamics. Similarly, crowd dynamics in public buildings such as museums have been studied by Hong et al. [13] using Wi-Fi traces of visitors. They worked to infer crowd trajectories from prior knowledge

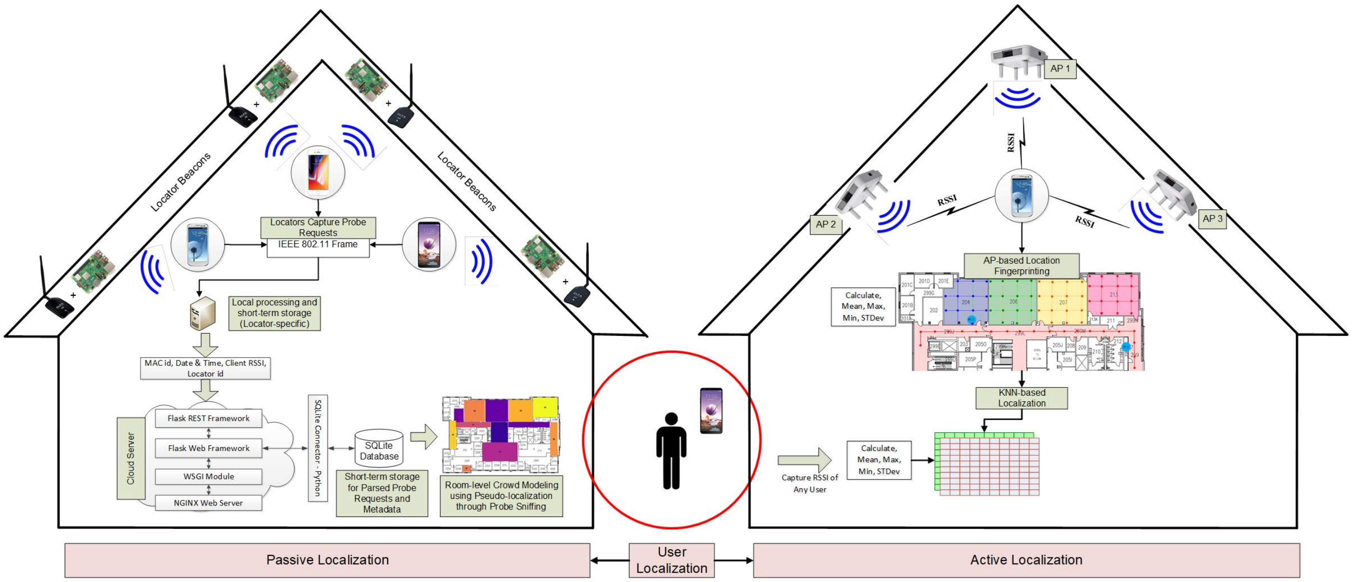


Figure 1: *MiamiMapper* System Architecture

rather than performing localization. Another work authored by Hu et al. [14], studied the energy consumption aspects of wireless APs and phones (for probe-based localization) both in small as well as large-scale environments.

In this paper we propose the *MiamiMapper* system which can capture the Wi-Fi probes requests of users and can localize individual users as well as capture flow of large crowds without requiring any Internet connection. The client smartphones just require to keep their Wi-Fi on in search of the neighboring APs which will emit probe packets. Our system captures those probe packets to uniquely identify the number of users and analyzes them for the crowd modeling. The entire system is developed using COTS hardware and software.

3 SYSTEM ARCHITECTURE

Figure 1 shows the system architecture of our proposed *MiamiMapper* system. The system is divided into active and passive localization parts.

Active localization is the approach by which pre-generated location fingerprints are used to localize a new user. Passive localization, on the other hand, uses some static Sensing Units (SSUs) w.r.to which a new user is localized. So, while the active scheme captures AP beacons to localize a user device, the passive scheme uses the probe requests emitted by a user device and captured by a set of Locator(s).

The system in overall has three different stakeholders – Client, AP and Locator. A Client is a person carrying a Wi-Fi enabled smartphone whose location is to be tracked. APs are the wireless access points available in the surrounding of the Client. The Locator side comprises of 4 Static Sensing Units (SSU) and a Cloud Server. SSUs are deployed at strategic points. Clients moving through the range of a SSU will be identified uniquely through their MAC ids. Below we describe the Client and Locator modules in more details.

3.1 Client Side

Client devices are wireless devices like laptops and smartphones having Wi-Fi support. We have tested with various smartphones running Android Operating Systems.

3.2 Locator Side

Locator side is divided into two parts – four Static Sensing Units (SSUs) and the Cloud Server.

3.2.1 Static Sensing Units (SSUs). Locator module consists of four Static Sensing Units (SSUs) each constructed using a Raspberry-Pi microcomputer connected to an Alfa [3] 150 Mbps Wireless (802.11n) Network Adapter which helps us to carry out passive scanning of all visible Wi-Fi devices in the monitor mode (Figure 6 (a)). The range of each Alfa adapter is 50-60 m (in open space) and 20-25 m (indoors). The task of SSU is to capture the probe requests transmitted by wireless devices within the wireless sensing range (to find an Access Point (AP) to associate with), process the probe requests to get their MAC ids (which remains un-encrypted even for a secure network), store them locally, and finally upload the information to the Cloud server once every 10 seconds. The probe-requests are captured in real time and processed. Multiple threads are spawned as needed to process the received packets in parallel. All the SSUs are connected through Wi-Fi interfaces with the same WLAN. PSU collects data (record) in the following format *<date and time, MAC-address, Client RSSI>* where, *MAC-id* is the MAC address of the Client device (*D*) detected, *Time* is the physical date and time at which *D* is detected, and *Client RSSI* is the signal strength of D to estimate the physical distance (in meters) between the SSU and *D*.

The detailed implementation and processing of SSU is explained in Section V Prototype Implementation.

3.2.2 Cloud Server. Back-end data storage consists of two connected entities: a Cloud server and a workstation. The cloud server

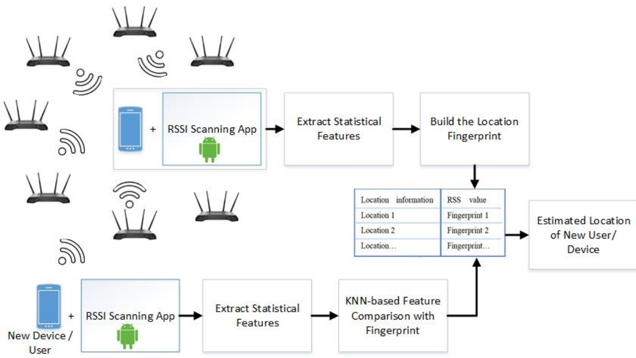


Figure 2: Block Diagram for Active Localization

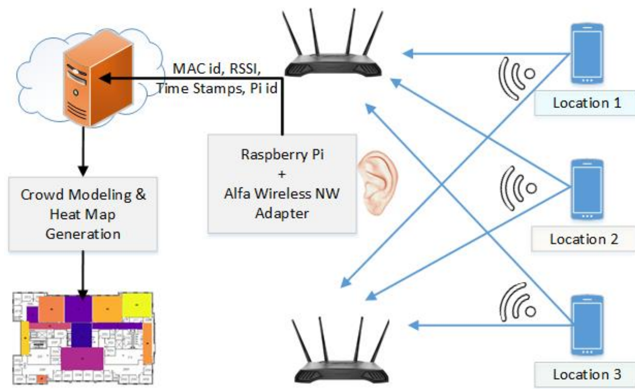


Figure 3: Block Diagram for Passive Localization

carries out the processing of probe requests captured and transmitted by the SSUs and uses the information for user localization by executing the algorithms associated with it. We use the Nginx Web Server for this purpose. We also have a workstation running Windows which has a scheduled task that runs every day at 10 AM to pull the aggregate data collected by all the 4 SSUs into the workstation to store it locally for the long term. The script sends a get request to an end-point in the Web Server daily at the scheduled time. When the server receives the request, it processes the 24-hour-data since the last request and compiles them into a single file and sends back to the Workstation.

4 ACTIVE AND PASSIVE LOCALIZATION

As already stated in Section 3, our proposed *MiamiMapper* system adopts both active and passive localization approaches. Active localization is achieved by measuring the signal strength (RSSI) from the AP beacons as observed by a client device and by building a location fingerprint. Passive localization is achieved by measuring the signal strengths from client devices (by capturing client probe requests) to the SSUs. In this section, we shall discuss in details the techniques adopted for active and passive localization in *MiamiMapper*.

4.1 Active Localization

Active Localization works by building the location fingerprints using one or more Client devices (smartphones) and then using those fingerprints to locate other Client devices.

4.1.1 Location Fingerprinting. The smartphones we used to build the fingerprints use an Android application (named *RSSI-tracker*) to scan the RSSI values of all the "visible" access points (APs) inside a building. This operation is called *location fingerprinting* and is a standard and widely used technique. We have used K-nearest Neighbor (K-NN) approach to pin-point user locations with respect to the nearby APs (see Figure 2). Assuming that the AP locations do not change and the environment remains largely static, we have developed the initial fingerprints using our Android app (*RSSI-tracker*). The carpet area of a floor is divided into square grids of size 10x10 feet. Then one user stands at each of the grid vertices and takes a slow 360 degree turn at the point spanning over 60 seconds holding the device in hand. The *RSSI-tracker* application (Figure 6 (c)) records the RSSI values of each "visible" APs at the frequency of 1/10 seconds and stores them in formatted files. A Python script is used to extract statistical features of the files, such as, MIN, MAX, MEAN and STDEV of RSSI. These statistical features are used to build a Wi-Fi based location fingerprint. For any new user we can estimate their location using this fingerprint. The user device requires a way to capture the RSSI values of nearby APs. Then we extract the statistical values of the RSSI observed and match those features with the previously generated fingerprint using KNN. The user location is approximated with an error of approximately 22–26 feet (or 6-8 meters).

4.2 Passive Localization

Passive localization is the technique opposite to active localization. It aims to passively monitor the probe requests generated from the Client devices through the SSU modules and then generate an estimate of user location plotted as a heatmap on the building's indoor plan.

We use the *RSSI-tracker* application in the user devices to emit probe requests which are then captured by the SSUs stationed around the building. The RSSI values, in this case, are the signal strengths measured from the user device to the SSUs. This process of probe monitoring is used for a 3-minute-interval for each room and then the generated data is labelled at each SSU with the corresponding room number. Using this, we build a room-level fingerprint of the RSSI statistical features which is mostly accurate.

When we localize a user, we compare the aggregate of their recently captured probe requests and find the closest matching room-level fingerprint, through a minimum mean-squared difference between a user's RSSI to each SSU compared to the known average RSSI fingerprint. We assume the user is located in that region where from the known average fingerprint was recorded.

We then repeat this process of localizing a user for every unique device (MAC id) seen in the past 30 minute-interval, which are overlapping. We sum up the number of users roughly localized in a room in that interval and pass it to heatmap for plotting on the map of the building (see Figure 3). So, we could actually track the trajectory of a user's movement over time (which we have not reported in this paper). 30-minute interval is chosen to trade off

(a) Benton Hall Exterior (b) First Floor Interior (c) Second Floor Classroom (c) Auditorium

Figure 4: Benton Hall, Miami University, Oxford, Ohio, USA [Source: https://www.thirdshiftphotos.com/]

between the number of probe requests generated and the accuracy of localization (if it's too long, the user might have moved to a far away or different location. If it's too short, there are not enough probe requests to confidently localize the user).

Whenever a MAC id is reported for more than 5 times, it is processed for localization. This is to avoid localizing accidental occurrings or one-time visitors. The localization algorithm runs constantly on the reported MAC ids and a heatmap is generated / refreshed every 10 seconds depending on the number of occupants reported / detected in a particular room. Since, the localization is accurate up to the room level, we have randomly distributed the occupant locations inside a room boundary in order to properly distribute the locations in the generated heatmap. We are using *simpleheat()* library for building the heatmap.

5 PROTOTYPE IMPLEMENTATION

This section depicts detailed implementation of the *MiamiMapper* system using COTS hardware and software modules. The *MiamiMapper* system has been developed and deployed for the indoor localization of users inside the Benton Hall in Miami University, Oxford, Ohio which is a 3-storey building (Figure 4) with total floor area of 82,661 sq ft (7,679.5 m2) and has seven different entrances/exits . Below we describe the system implementation in detail.

5.1 Active Localization using *MiamiMapper*

The active localization using the *MiamiMapper* system works by building the location fingerprint using the approach described in Section 4.1. We have a map with the locations of all the wireless APs inside the Benton Hall which has been used to build the location fingerprint.

5.2 Passive Localization using *MiamiMapper*

Our passive localization approach requires the setting up of the SSUs and the Web Server and then using those to monitor the probe packets generated by the users. Below we describe the procedures in details.

5.2.1 SSU Operation. The four SSU's used to develop the *MiamiMapper* system are statically deployed at strategic locations inside the Benton Hall in order to cover the entire building including all the 7 entrances/exits. They perform passive scanning in monitor mode where it just listens to the 802.11 channel and intercepts probe requests generated by Client devices.

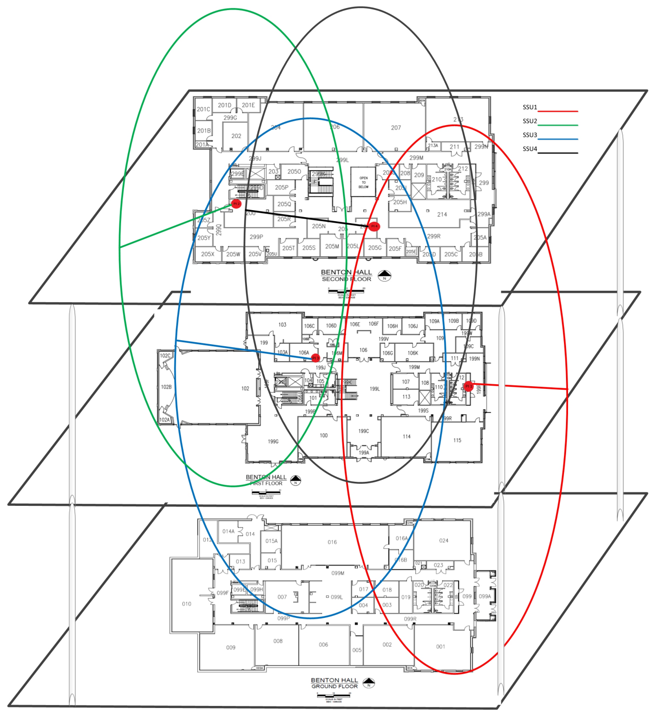


Figure 5: Range and Distribution of SSUs in Benton Hall [Not in scale]

The SSUs use Kali Linux as their Operating Systems because it comes with necessary tools and libraries required for the probe request capturing and processing. The code is based on an existing project repository called Probemon developed by Nik Harris [11]. All of the SSUs are running open SSH to make sure that we can access them from anywhere and at anytime since they are otherwise not accessible over the network. We have the SSUs check-in through a server interface which allows users to quickly get the connection information and the real-time status of the SSUs. This helps us to detect the health of the SSUs when they stop reporting while they still may be sniffing the network (just the networking capability is affected).

The data uploading task is handled by a separate thread in each SSU, which uploads probe-scans in a JSON format. After uploading a file successfully, it is still stored in the internal storage of the SSU for one whole month from the date of scan. If uploading fails, it retries to upload the file when network connection is available again. The file is uploaded using the Wi-Fi connection.

Together the SSUs send data to the cloud server every 10 seconds with a total average cloud access of 25000 times in 24 hours (for all the 4 SSUs). Each such request contain an average of 20 captured probe packets. This is an average of 500,000 probe requests captured per day.

5.2.2 Server Operation. Cloud Server module in Figure 1 (passive localization part) shows the interaction between various components on the Cloud. We used Flask micro-framework which uploads the raw data in the SQLite3 database. The database keeps track of the count of unique MAC ids seen. If a MAC id is seen for 5 or more times in total, then it is used for further processing. Below 5 citations, a MAC id is considered unusual and not reported. SSUs upload all detected MAC records (except the APs) which helps us for traffic analysis, estimation of congregation strength, crowd evacuation analysis, etc. Therefore, all MAC records will be stored on the Cloud Server in a time-stamped manner and are used for Client localization.

5.2.3 Passive Localization. We have used two different user devices – Samsung Galaxy S7 Edge and Samsung Galaxy S8 for building the room level fingerprints following the procedures described in Section IV B. Using two different phones gives us greater variety of RSSI values, since each wireless adapter has different capabilities. Fig. 6 depicts the range and distribution of the SSUs deployed in the Benton Hall. Since, all the rooms are not covered by all the SSUs we generate room-level fingerprint in a SSU-specific manner. E.g., from room 214, average RSSI to SSU4 is -67 dB whereas average RSSI to SSU2 is -89 dB.

6 PERFORMANCE EVALUATION AND RESULTS

The *MiamiMapper* system has been developed and extensively tested for indoor localization in the Benton Hall of Miami University. In this section, we shall introduce our experiment results followed by the performance metrics used to test and evaluate the system. Our system has started working and data logging from March 18, 2019 and continues since then. The data stored in the served is used to analyze the performance using the following metrics.

6.1 Performance Metric

We tested the *MiamiMapper* system using the following performance metrics.

- No. of Unique Devices (N_d): It is the count of the number of unique MAC ids seen through a day or week. This is the prime metric for crowd measurement.
- No. of Processed Packets: It is the measure of the total number of probe packets being handled by the SSUs in an hourly basis.
- Location accuracy (L_{acc}): It is the distance (measure in meters) between the actual location of the user w.r.to the predicted location. Location accuracy has been calculated for both active and passive localization schemes.

6.2 Results of Crowd Analysis

Initially we have measured the number of unique devices varying through a typical working week and plotted it in the (Figure 7 (a)). We can observed that the number of devices increases during the mid-day of the working days and gradually reduces during the nights and the weekends. Similar scenarios are repeated during another week (March 18-24, 2019) before the Annual Spring Break (Figure 7 (b)). However, we can notice that the weekends are almost flat as the students left early for the break and the trend continues for the whole of next week (Spring Break week – March 25-31, 2019) as shown in (Figure 7 (c)). A final trend of crowds has been shown in (Figure 7 (d)) during the pre-exam week (April 29 – May 05, 2019) with higher activities during the weekend as the students congregate in the labs to finish their last moment studies and assignments. One important issue to note both in Figure 7 (b) and (c) is that the SSU3 has been shut down to test the scalability of the system (see following paragraph for more details). We can observe that the Figure 7 (b) and (d) record same/similar number of unique MAC ids observed which shows that the system is capable of handling the load even with one SSU down at any time.

A closer look of some of the days (Figure 8 (a) and (b)) shows that each of the SSUs (Raspberry Pi's) (Figure 6 (a)) actively capture probe packets except SSU3. We deliberately shut SSU3 to check whether the remaining 3 SSUs can handle the load while still covering all the users inside the building. Similarly we tested by shutting each of the other SSUs (only one at a time). We can compare results of Wednesday 03/20 in (Figure 8 (a)) with Monday 04/22 in (Figure 8 (b)) and can observe that the results are almost similar although the former has only 3 SSUs working unlike the latter in which all the 4 SSUs are working. Moreover, we can see that the number of unique devices varies throughout the day with highest number of variations observed 10 AM to 4 PM which is our typical class times. Also, we have noticed that the number of probe packets increase when students move around (go from one class to the next) rather than sitting in a classroom during lecture. When the smartphones are not in active use, the number of probe packets fall very significantly although it further depends on the type of the wireless adapter.

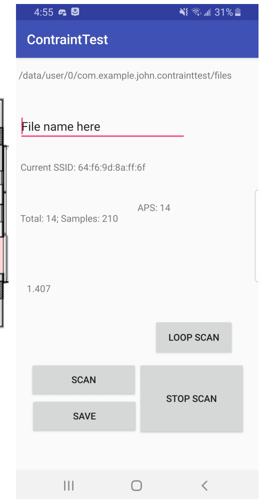
Figure 9 shows a cumulative number of packets being processed by the four SSUs on an hourly basis for a week from April 27 to May 3, 2019. We can clearly see that the trends are similar for each of the SSUs. This shows that the SSUs are very well distributed across the building and they sense almost same number of probe packets including the duplicate ones.

6.3 Performance of Active Localization

Our approach for active localization builds the location fingerprint by measuring the RSSI values observed by the client devices from different APs around the different rooms in the second floor of the Benton Hall. Three sets of data values are measured w.r.to different phones and different times of the day/week and the observed RSSI values are presented in (Figure 10). This is because each phone has a different wireless adapter and the RSSI values vary with number of people in the building. The three sets of data collected for building the location fingerprints are as below.

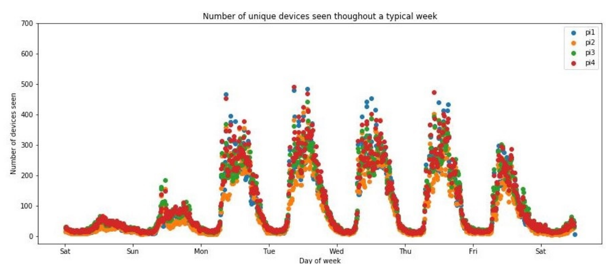
(a) Static Sensing Unit (SSU)

(b) Second Floor Map of Benton Hall

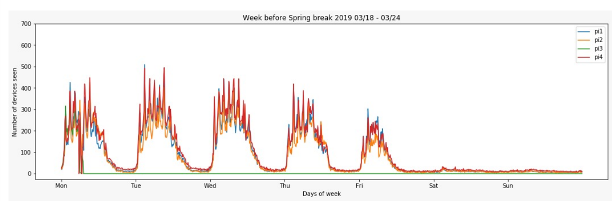


(c) *RSSI-Tracker* Android App

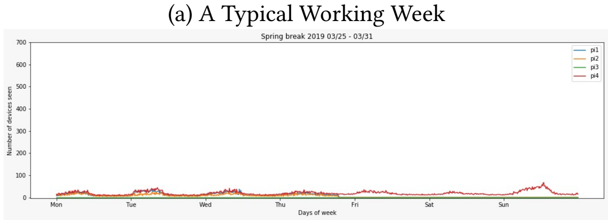
Figure 6: Hardware and Software Used in *MiamiMapper* System



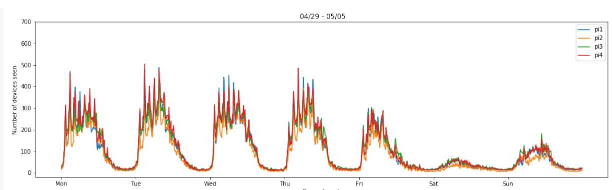
(a) A Typical Working Week



(b) Week Before the Spring Break



(c) The Spring Break Week



(d) Week Before the Final Exam

Figure 7: No. of Unique Devices (N_d) Observed During Various Weeks

Set 1: Measured Using Samsung Galaxy S8 on Oct. 19, 2018 [Friday, 12:05 – 12:31 PM]

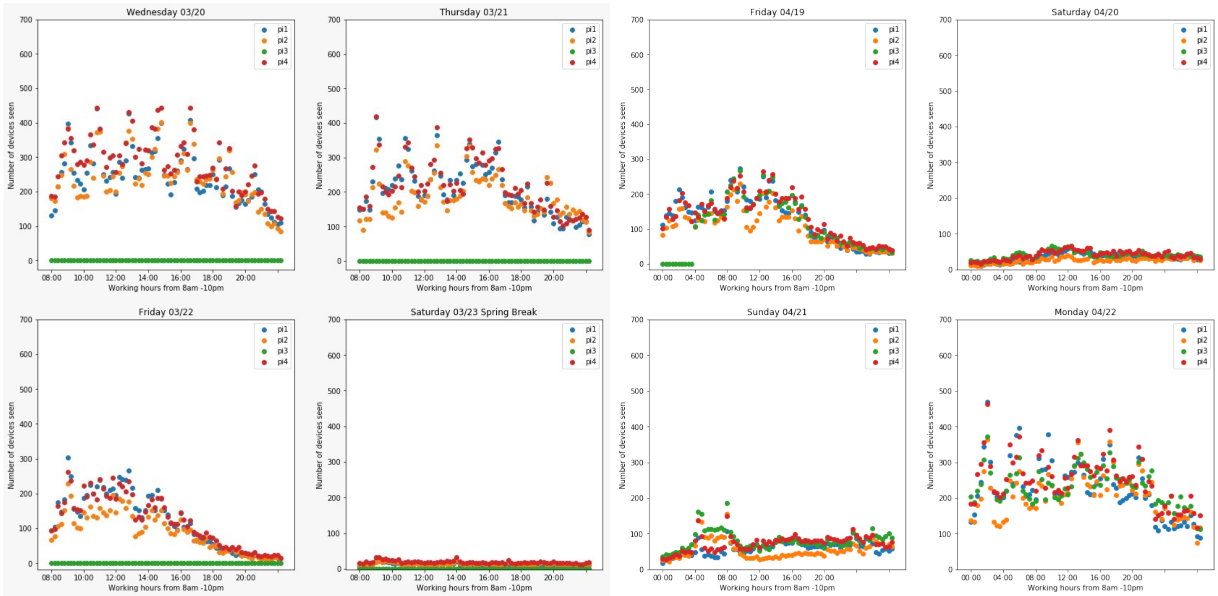
Set 2: Measured Using Samsung Galaxy S7 Edge on Jan. 19, 2019 [Saturday 4:55-5:37 PM]

Set 3: Measured Using Samsung Galaxy S7 Edge on Jan. 20, 2019 [Sunday 5:18-5:36 PM].

Set 2 and Set 3 are collected on the weekend evenings to ensure minimum absorption of wireless signals by human body. They give us the benchmark values with minimum building occupancy provided no other environmental conditions change. Since, Set 1 was collected on a weekday we can clearly see in Figure 10 how the RSSI values widely fluctuate for Set 1 compared to the other two.

With respect to the aforementioned 3 datasets, we have calculated the accuracy of localization (L_{acc}) using the cross-validation technique. Any two datasets are used for training while the remaining one is used for testing. We can observe that the average accuracy is 7.28 meters which is significantly well (Figure 11 (a)) compared to the state-of-the-art.

We have further calculated the false positives, false negatives, true positives and true negatives of location prediction using our generated location fingerprints. The confusion matrix shown in (Figure 11 (b)). describes the accuracy of our fingerprinting algorithm. The values range between 0-3 which represents the number of guesses / predictions for each location. White squares represent 3 out of 3 guesses are correct (true positive), whereas yellow one represents 2 guesses out of 3 are correct and red represents 1 guess out of 3 is correct. We can notice that in majority of the case, 2-3



(a) Four Days in March (March 20-23, 2019) with SSU3 Turned Off

(b) Four Days in April (April 19-22, 2019)

Figure 8: No. of Unique Devices (N_d) Observed During Various Days

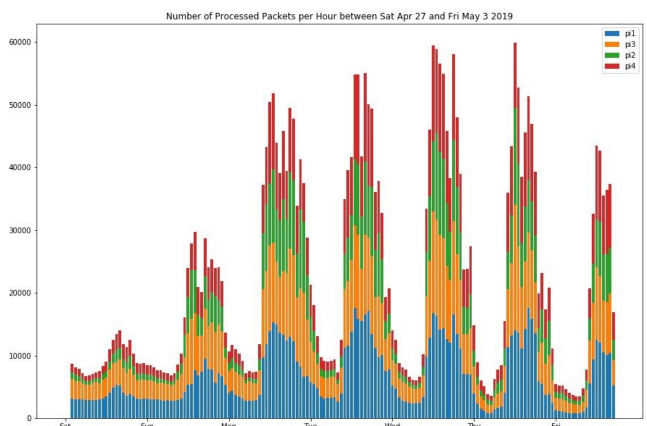


Figure 9: No. of Processed Packets Per Hour Between April 27-May 03, 2019

guesses are correctly localizing the user. This shows the efficiency of our location fingerprinting technique used for active localization.

In order to establish the efficiency of the localization approaches used in *MiamiMapper*, we have compared the location accuracies of active and passive localization techniques with the location accuracy of GPS based approach in indoor environment (Figure 11 (c)). Since, GPS is known to perform poorly in indoor environment, it has a poor accuracy of 30 m on average over the 50 samples taken at different locations inside the Benton Hall spread across the three floors. The location accuracy of passive localization approach using MiamiMappper is obtained to be 7 meters from the heatmaps reported in Figure 12 (see Section 6.4).

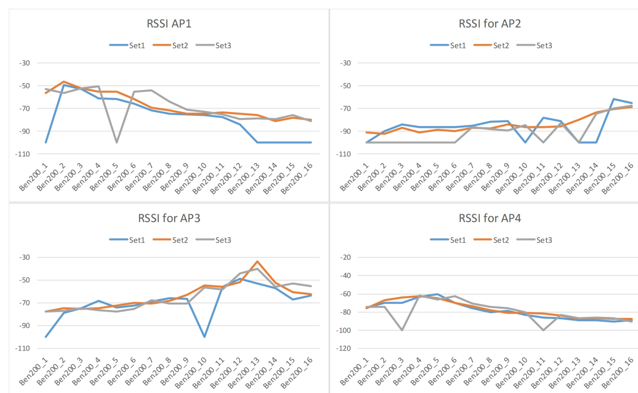
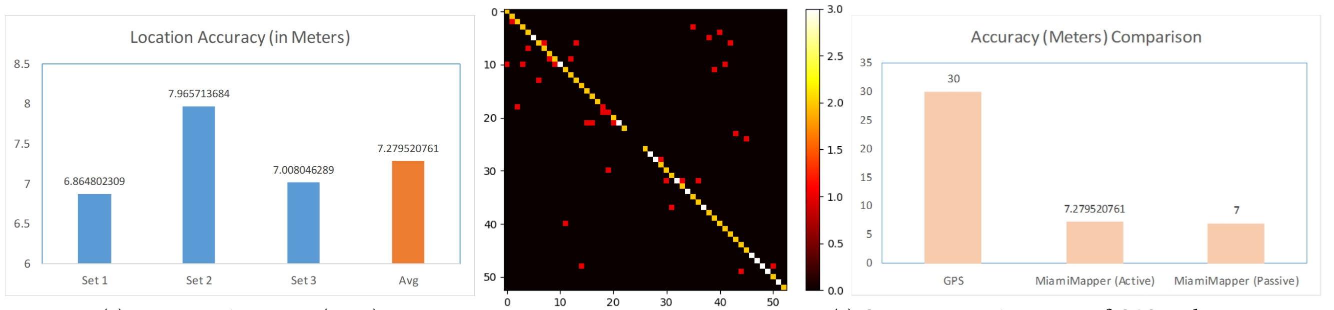


Figure 10: Variation of RSSI values for Active Localization

6.4 Performance of Passive Localization

Using our passive localization approach described in Section 4.2, we have estimated the location of users present inside the Benton Hall and have plotted them using a heatmap which shows the occupancy of different rooms inside the building (Figure 12 (a) to (f)). Our passive localization approach actually detects and localizes crowds (combining locations of multiple individual users) and it works by generating a room level fingerprint. Thus, the accuracy of passive localization is also measured w.r.to the fingerprinting accuracy. Through trial and error type experimental methods we have decided that the accuracy is close to 7 meters. This is supported by the fact that the heatmaps presented in Figure 12 (b) gives the best crowd estimate in respective rooms of the building. With lower radius (4 meters) (Figure 12 (a) and (e)) the individual user locations are not quite correctly guessed and the crowd formations are not detected.



(a) Location Accuracy (L_{acc}) (b) Confusion Matrix of (L_{acc}) (c) Comparative Accuracy of GPS and *MiamiMapper*

Figure 11: Accuracy of Active Localization Using *MiamiMapper* System

Similarly, with higher radius (11 or 15 meters), we can easily observe (Figure 12 (c), (d) and (e)) that the occupants in different rooms (classes going on in neighboring rooms) are combined together and considered as a large crowd.

The results are continuously updated in real time and the heatmap gets refreshed every 10 seconds. The heatmap at current instant can be observed through the domain *www.miamimapper.com.*

7 CONCLUSION AND FUTURE WORKS

Monitoring large crowds as well as modelling and analyzing their movement patterns and dynamics is non-trivial. Research in this domain is mainly based on image or video analysis which is unsuitable for real-time situational analysis for quick decision making. In this paper, we have developed an indoor crowd analysis system, called *MiamiMapper*, for a university building with high occupancy during working hours. Our system plans to passively listen to the probe packets being generated by wireless devices of users while they are trying to associate with the wireless APs inside the building. Even if a wireless device is already connected to an AP, it will try to find a better AP, so it will still emit probe requests but rather infrequently. Since, the probe packets contain the unique MAC id of the devices, it is possible to keep a count of the unique devices present inside the building along with their trajectories over time. The *MiamiMapper* system localizes users (with very high accuracy) and thereby the crowds and analyzes their movement patterns using a continuously changing heatmap of the floorplan. We have further analyzed and compared the localization accuracy of users through active fingerprint based localization and passive probe monitoring based one.

In future, we plan to build a mathematical model of the crowd movement pattern and study it in more detail. We also want to test with iOS devices and various other types of wireless adapters. Some of our other challenges include identifying the devices connecting from outside the building and improving the location accuracy further. Addressing the general concern about privacy is also our foremost future concern.

8 ACKNOWLEDGMENTS

We acknowledge Department of Computer Science and Software Engineering, Miami University (MU), OH, USA for their partial financial support for deploying this system. We also thank Dr. Manar Mohammad (MU), Dr. Utkarsh Singh (ULB, Brussels) and Ms Anjali Purayil (UG student at MU) for their insightful comments and technical help. This work has also been partially supported by the Miami University CFR Summer Research Award 2019 to Dr. Vaskar Raychoudhury.

REFERENCES

[1] Yuvraj Agarwal, Bharathan Balaji, Rajesh Gupta, Jacob Lyles, Michael Wei, and Thomas Weng. 2010. Occupancy-driven Energy Management for Smart Building Automation. In *Proceedings of the 2Nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building (BuildSys '10)*. ACM, New York, NY, USA, 1–6. https://doi.org/10.1145/1878431.1878433

[2] A. R. Al-Ali, F. A. Aloul, N. R. Aji, A. A. Al-Zarouni, and N. H. Fakhro. 2008. Mobile RFID Tracking System. In *2008 3rd International Conference on Information and Communication Technologies: From Theory to Applications*. 1–4. https://doi.org/10.1109/ICTTA.2008.4530117

[3] Alfa.com. 2019. Alfa Wireless Adapter (AWUS036H), Wireless-G with SMA Adapter (2dBi or 5dBi Antenna). http://www.alfa.com.tw/products_show.php?pc=34&ps=92. (2019).

[4] Marco V Barbera, Alessandro Epasto, Alessandro Mei, Vasile C Perta, and Julinda Stefa. 2013. Signals from the crowd: uncovering social relationships through smartphone probes. In *Proceedings of the 2013 conference on Internet measurement conference*. ACM, 265–276.

[5] F. Calabrese, M. Colonna, P. Lovisolo, D. Parata, and C. Ratti. 2011. Real-Time Urban Monitoring Using Cell Phones: A Case Study in Rome. *IEEE Transactions on Intelligent Transportation Systems* 12, 1 (March 2011), 141–151. https://doi.org/10.1109/TITS.2010.2074196

[6] G. De Angelis, G. Baruffa, and S. Cacopardi. 2013. GNSS/Cellular Hybrid Positioning System for Mobile Users in Urban Scenarios. *IEEE Transactions on Intelligent Transportation Systems* 14, 1 (March 2013), 313–321. https://doi.org/10.1109/TITS.2012.2215855

[7] P. C. Deepesh, Rashmita Rath, Akshay Tiwary, Vikram N. Rao, and N. Kanakalata. 2016. Experiences with Using iBeacons for Indoor Positioning. In *Proceedings of the 9th India Software Engineering Conference (ISEC '16)*. ACM, New York, NY, USA, 184–189. https://doi.org/10.1145/2856636.2856654

[8] P. Fuxjaeger, S. Ruehrup, T. Paulin, and B. Rainer. 2016. Towards Privacy-Preserving Wi-Fi Monitoring for Road Traffic Analysis. *IEEE Intelligent Transportation Systems Magazine* 8, 3 (Fall 2016), 63–74. https://doi.org/10.1109/MITS.2016.2573341

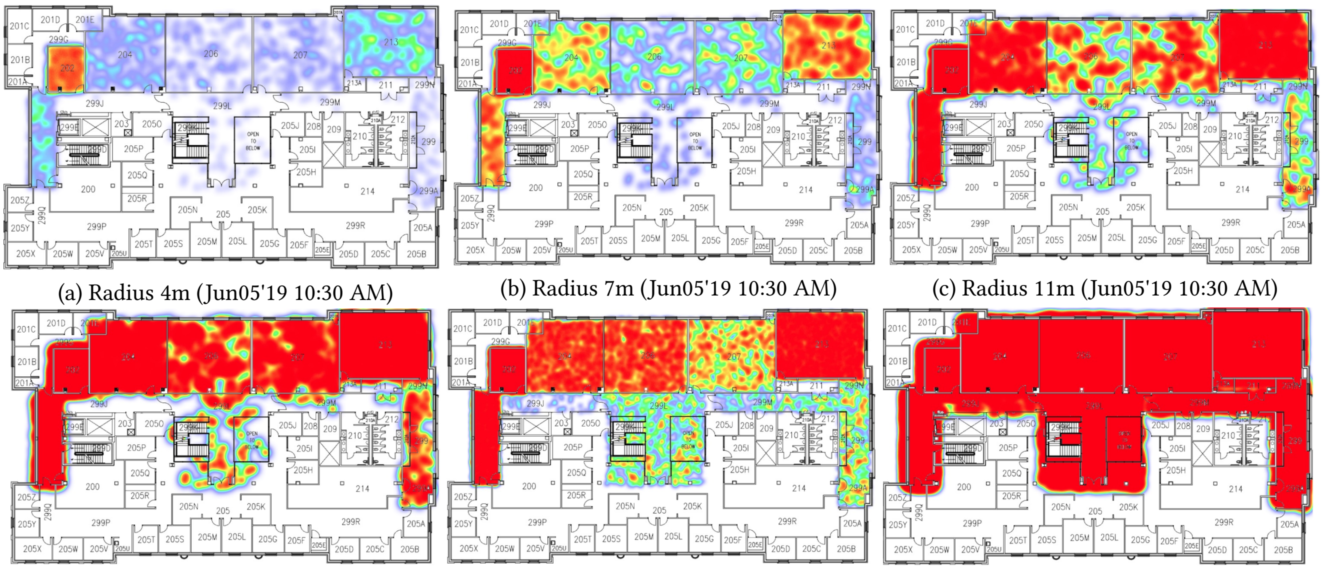
[9] Shaogang Gong, Chen Change Loy, and Tao Xiang. 2011. *Security and Surveillance*. Springer London, London, 455–472. https://doi.org/10.1007/978-0-85729-997-0_23

[10] Y. Gu, A. Lo, and I. Niemegeers. 2009. A survey of indoor positioning systems for wireless personal networks. *IEEE Communications Surveys Tutorials* 11, 1 (First 2009), 13–32. https://doi.org/10.1109/SURV.2009.090103

[11] Nik Harris. 2019. Probemon. https://github.com/nikharris0/probemon. (2019).

[12] S. He and S. . G. Chan. 2016. Wi-Fi Fingerprint-Based Indoor Positioning: Recent Advances and Comparisons. *IEEE Communications Surveys Tutorials* 18, 1 (Firstquarter 2016), 466–490. https://doi.org/10.1109/COMST.2015.2464084

[13] Hande Hong, Girisha Durrel De Silva, and Mun Choon Chan. 2018. CrowdProbe: Non-invasive Crowd Monitoring with Wi-Fi Probe. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 3, Article 115 (Sept. 2018), 23 pages. https://doi.org/10.1145/3264925



(a) Radius 4m (Jun05'19 10:30 AM) (b) Radius 7m (Jun05'19 10:30 AM) (c) Radius 11m (Jun05'19 10:30 AM)

(d) Radius 15m (Jun05'19 10:30 AM) (e) Radius radius 4m (May30'19 10:30 AM) (f) Radius 15m (May30'19 10:30 AM)

Figure 12: Heatmap for Passive Localization Using *MiamiMapper* System

[14] Xueheng Hu, Lixing Song, Dirk Van Bruggen, and Aaron Striegel. 2015. Is there wifi yet?: How aggressive probe requests deteriorate energy and throughput. In *Proceedings of the 2015 Internet Measurement Conference*. ACM, 317–323.

[15] Taeyu Im and Pradipta De. 2016. User-Assisted OCR on Outdoor Images for Approximate Positioning. In *Information Science and Applications (ICISA) 2016*, Kuinam J. Kim and Nikolai Joukov (Eds.). Springer Singapore, Singapore, 1419–1429.

[16] Cisco Inc. 2019. Cisco Wireless Location Appliance. http://www.cisco.com/c/en/us/products/collateral/wireless/wireless-location-appliance/product_data_sheet0900aecd80293728.html. (2019).

[17] Tarun Kulshrestha, Divya Saxena, Rajdeep Niyogi, Vaskar Raychoudhury, and Manoj Misra. 2017. SmartITS: Smartphone-based identification and tracking using seamless indoor-outdoor localization. *Journal of Network and Computer Applications* 98 (2017), 97 – 113. https://doi.org/10.1016/j.jnca.2017.09.003

[18] Libelium. 2019. Libelium, Smartphone Detection. http://www.libelium.com/products/meshlium/Smartphone-detection/. (2019).

[19] H. Liu, H. Darabi, P. Banerjee, and J. Liu. 2007. Survey of Wireless Indoor Positioning Techniques and Systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 37, 6 (Nov 2007), 1067–1080. https://doi.org/10.1109/TSMCC.2007.905750

[20] Halgurd S Maghdid, Ihsan Alshahib Lami, Kayhan Zrar Ghafoor, and Jaime Lloret. 2016. Seamless outdoors-indoors localization solutions on smartphones: implementation and challenges. *ACM Computing Surveys (CSUR)* 48, 4 (2016), 53.

[21] T. Mantoro, A. D. Jaafar, M. F. M. Aris, and M. A. Ayu. 2011. HajjLocator: A Hajj pilgrimage tracking framework in crowded ubiquitous environment. In *2011 International Conference on Multimedia Computing and Systems*. 1–6. https://doi.org/10.1109/ICMCS.2011.5945629

[22] R. O. Mitchell, H. Rashid, F. Dawood, and A. AlKhalidi. 2013. Hajj crowd management and navigation system: People tracking and location based services via integrated mobile and RFID systems. In *2013 International Conference on Computer Applications Technology (ICCAT)*. 1–7. https://doi.org/10.1109/ICCAT.2013.6522008

[23] M. Mohandes. 2011. Pilgrim tracking and identification using the mobile phone. In *2011 IEEE 15th International Symposium on Consumer Electronics (ISCE)*. 196–199. https://doi.org/10.1109/ISCE.2011.5973812

[24] L. M. Ni, D. Zhang, and M. R. Souryal. 2011. RFID-based localization and tracking technologies. *IEEE Wireless Communications* 18, 2 (April 2011), 45–51. https://doi.org/10.1109/MWC.2011.5751295

[25] Janice Partyka. 2012. A Look at Small Indoor Location Competitors. https://www.gpsworld.com/wirelesslook-small-indoor-location-competitors-13229/. (2012).

[26] W. Pattanusorn, I. Nilkhamhang, S. Kittipiyakul, K. Ekkachai, and A. Takahashi. 2016. Passenger estimation system using Wi-Fi probe request. In *2016 7th International Conference of Information and Communication Technology for Embedded Systems (IC-ICTES)*. 67–72. https://doi.org/10.1109/ICTEmSys.2016.7467124

[27] V. Raychoudhury, S. Shrivastav, S. S. Sandha, and J. Cao. 2015. CROWD-PAN-360: Crowdsourcing Based Context-Aware Panoramic Map Generation for Smartphone Users. *IEEE Transactions on Parallel and Distributed Systems* 26, 8 (Aug 2015), 2208–2219. https://doi.org/10.1109/TPDS.2014.2345067

[28] Liqing Ren, Xiaojiang Chen, Binbin Xie, Zhanyong Tang, Tianzhang Xing, Chen Liu, Weike Nie, and Dingyi Fang. 2016. DE2: localization based on the rotating RSS using a single beacon. *Wireless Networks* 22, 2 (01 Feb 2016), 703–721. https://doi.org/10.1007/s11276-015-0998-9

[29] J. C. Silveira Jacques Junior, S. R. Musse, and C. R. Jung. 2010. Crowd Analysis Using Computer Vision Techniques. *IEEE Signal Processing Magazine* 27, 5 (Sep. 2010), 66–77. https://doi.org/10.1109/MSP.2010.937394

[30] Mathias Versichele, Tijs Neutens, Matthias Delafontaine, and Nico Van de Weghe. 2012. The use of Bluetooth for analysing spatiotemporal dynamics of human movement at mass events: A case study of the Ghent Festivities. *Applied Geography* 32, 2 (2012), 208 – 220. https://doi.org/10.1016/j.apgeog.2011.05.011

[31] Wei Wang, Raj Joshi, Aditya Kulkarni, Wai Kay Leong, and Ben Leong. 2013. Feasibility Study of Mobile Phone WiFi Detection in Aerial Search and Rescue Operations. In *Proceedings of the 4th Asia-Pacific Workshop on Systems (APSys '13)*. ACM, New York, NY, USA, Article 7, 6 pages. https://doi.org/10.1145/2500727.2500729

[32] M. Wirz, T. Franke, D. Roggen, E. Mitleton-Kelly, P. Lukowicz, and G. Tröster. 2012. Inferring Crowd Conditions from Pedestrians' Location Traces for Real-Time Crowd Monitoring during City-Scale Mass Gatherings. In *2012 IEEE 21st International Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises*. 367–372. https://doi.org/10.1109/WETICE.2012.26

[33] Jiang Xiao, Zimu Zhou, Youwen Yi, and Lionel M. Ni. 2016. A Survey on Wireless Indoor Localization from the Device Perspective. *ACM Comput. Surv.* 49, 2, Article 25 (June 2016), 31 pages. https://doi.org/10.1145/2933232

[34] C. Yang and H. Shao. 2015. WiFi-based indoor positioning. *IEEE Communications Magazine* 53, 3 (March 2015), 150–157. https://doi.org/10.1109/MCOM.2015.7060497