

A passive, non-intrusive, cheap method to identify behaviours and habits in the Campus

Javier Andión, José M. Navarro, Manuel Álvarez-Campana, Juan C. Dueñas
Departamento de Ingeniería de Sistemas Telemáticos,

Universidad Politécnica de Madrid

ETSI Telecomunicación, av. Complutense 30, 28040 Madrid

{j.andion,josemanuel.navarro,manuel.alvarez-campana,juancarlos.duenas}@upm.es

Abstract- Infrastructure usage discovery, positioning and analysis of behaviours of its users usually requires a collection of accurate and frequent positioning data. This paper shows how a network of inexpensive and non-intrusive sensors can serve to perform this kind of analysis by detecting devices with Wi-Fi connectivity.

By this analysis, we show that, although individual tracking is not possible because of limitations of sensors, we can obtain the hours of use of the infrastructures, the occupation of the different areas at each moment and some of the most common users' behaviours.

Keywords- Wi-Fi sensors, MAC address, semantic locations, behaviour patterns discovery

I. INTRODUCTION

Knowing the way users of a public infrastructure behave is a key to allocate the resources that must be assigned, ensure its safety, and control the usage. In this paper we propose a method, based on a minimal communication and computing infrastructure, to discover users individual and collective behaviour as regards usage of buildings.

Indoor location or pedestrian location has been a key research topic in last years [1]. Most works aim to discover the fine grain movements of people inside buildings, by using the mobile network, or personal area networks; these systems try to help the users to discover the path in a building, measure the length of stay in a mall for commercial purposes, or simply to ease people movements by removing obstacles, aside allowing for automatic movements of objects [2]. These works are based on location methods that make it possible to discover detailed paths in buildings, but

they require either installing Bluetooth beacons or the cooperation of the mobile network antennas [3].

Based on the description of the region of interest, the concept of semantic trajectory has emerged as a key element to relate people trajectory with the activities they perform on it. In order to reason about people's habits this concept conveys more information than the pure trajectory, since the trajectory (series of points and times a people moves) is enhanced with labels marked as points of interests; even though the trajectory may not be so granular, thus getting to the concept of regions of interest. So, the semantic trajectory evolves to become the series of regions of interests visited with the time elapsed in each. We state that it is this information what counts, for example, to identify behaviours [4], and not much information is required for such purposes. This information can be obtained easily in a passive, non-intrusive way by using Wi-Fi probes emitted by autonomic antennas, apart of the data network giving service to the users, and providing most of users carry a smartphone with Wi-Fi capabilities.

We have applied this method to the analysis of user behaviour in campus; this is a public installation that anybody can use -up to a certain extent-, on which it is difficult to cover all places using cameras, expensive to cover with PAN beacons, and at the same time it is very interesting to know the movements of groups of students, the behaviour of the staff, the usage of shared areas (library, rest rooms, cantina, meeting rooms), the average number of people using the facilities and their non-regular usage. Key elements in this approach are: the minimal non-intrusive infrastructure required, the

small amount of information handled and its inaccuracy, and the simplicity of algorithms used to discover patterns of behaviour -mostly queries on the dataset.

II. APPLICATION SCENARIO

Universidad Politécnica de Madrid (UPM), as part of its City of the Future initiative, deployed a platform for experimentation composed by more than one hundred sensors able to perform real-time monitoring of 20 buildings in its Campus de Excelencia Internacional de Moncloa, gathering information on power consumption, environmental parameters, light, and buildings occupancy. The platform also includes real-time storing, processing, analysis and visualization of data. During a normal day, there are usually around 4000 people in the school: around 3000 students, at most 500 teachers and researchers and less than 500 administrators and maintenance staff.

Understanding how space and the installations were used by the students and the staff at the university was soon proposed as one of the key insights extracted from the data, and therefore the need to detect presence of people was of paramount importance. The research group in charge opted to building sensors able to detect Wi-Fi devices, also known as Wi-Fi tracking [5]; this solution has already used to analyze usage of public transportation (London underground nov 2016) or movements in public spaces such as airports. Then, they decided to develop cheap sensors for Wi-Fi tracking, based on Raspberry Pi boards with an external Wi-Fi module able to perform passive monitoring (Fig. 1. shows a photograph). The sensors are connected to power avoiding the usage of batteries. They read the header of radio IEEE 802.11 packages in its region of reach, and extract the MAC addresses of devices. As these MAC addresses are unique per device, counting them is a good indicator of the number of devices available on the sensor's surroundings, and they allow for temporal correlation analysis, thus obtaining useful information such as stay time, availability patterns, temporal patterns. All in all, and as MAC addresses are considered personal data under Spanish law (Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal), the system anonymizes data for analysis by an irreversible hash MD-5 function with salt applied by the sensors, which avoids brute-force attacks with pre-computed tables. Once anonymized, data is moved to a central server in the University by means of Ethernet connection -all buildings are cabled this way.

We have been provided data from one set of buildings equipped with 9 sensors, strategically allocated in certain points of interest. By means of Wi-Fi probes, each sensor scans its surroundings each minute in all Wi-Fi channels, so most of the Wi-Fi devices are captured, anonymized and stored in the sensor. For

devices, we have checked that probe requests are sent between 20 seconds and 60 seconds period, depending on the type of smartphone. So, we were provided with log text files per sensor containing in each line: anonymized MAC address seen and its timestamp.



Fig. 1. Raspberry-pi Wi-Fi sensors

Our analysis is based on aggregating all the information in the log files of the sensors for a certain time span, counting identifiers in each sensor, identifiers in each time period, and then trajectories of each identifier, frequent trajectories, and behaviour profiles based on trajectories. One month renders: $n * 30 * 24 * 60 * ids$, where n is the number of sensors and ids is the average number of identifiers in our data. We processed them using the Spark processing libraries, whose execution speed eased the execution of many tests in this exploration exercise.

Specifically, we have used all the data collected during the whole month of 2016 May. This dataset is composed by 8.3M samples, where each sample is one register of one user seen by one sensor in one certain minute. Throughout the month 18K different devices are detected, this number reduces to 10K for devices seen more than 30 minutes in the month. On average 3.5K different devices are seen daily. These data were collected by 9 sensors:

1. Building A entrance (“Entr A”): this sensor is placed in the main entrance to the installations
2. Building A secondary entrance (“Entr A Sec”): this sensor is located in one of the secondary access to building A, it covers most of the classrooms of this building.
3. Work and study tables (“Std Tables”): this sensor is close to “Entr A”, just above an area of tables where students gather to work and study in groups.
4. Library (“Library”): the sensor is inside the library, which is open every day from 9:00 to 21:00.
5. Building B entrance (“Entr B”): building B contains student’s laboratories and offices.
6. Building B laboratories entrance (“Entr B Lab”): this sensor is placed in the secondary

entrance of this building closest to the laboratories.

7. Building B secondary entrance (“Entr B Sec”): this sensor is located in one of the secondary access to building B covering the most used classrooms of this building.
8. Building C entrance (“Entr C”): building C contains research laboratories and offices. Students do not have classes in this building.
9. Building D entrance (“Entr D”): this last sensor covers the entrance of building D and the library back.

Further on, Fig. 5 shows the positions of these sensors represented on the school map.

It is clear that results can only be approximations to the usage of space and infrastructure in the Campus, as there are many sources of error: not all people carry a mobile phone with its Wi-Fi capability turned on (students are suggested to switch off their mobile phones while at lectures or at library); on the contrary, some users carry more than one device; coverage area of sensors is severely conditioned by physical disposition of buildings; there are errors in capturing a mobile phone by the sensor; coverage areas overlap; sensors are put on a 2D map while devices move in a 2.5D space (at least two floors). But even under these limited conditions and applying simple algorithms we have been able to get a hint on people movements and behaviour, and to identify user types.

The main concepts in our study are defined now:

- Device position: we are not using power measurements, so the only valid approximation for the position of a single device in a given time is the point where the sensor is located. For devices seen by more than one sensor at once we have allocated to the sensor that saw for more timeslices.
- Regions or zones: we defined a zone per sensor, as the places where a device is detected by this sensor. Ultimately, if all sensors would get power enough the regions would define a Voronoi map, but as we do not know the effective reach of sensors we can only speculate on this.
- Timeslice: we are using 1-minute timeslices, as sensors are able to launch Wi-Fi probes at that pace. Later, for temporal analysis, we aggregate timeslices to create 1-hour sensing windows.
- Semantic location: sensors are located close to places where people actually do something (studying at the library, attending lectures at classrooms, performing experiments at labs, having lunch at the cantina, etc.). Particularly useful are sensors located in entry/exit places. So, once we know the sensor zone a device is

in, we can infer the most likely location and annotate this with the most likely activity. In fact, sensors names reflect this: Library, Entr B Lab, Entr A Sec, Entr D, Entr A, Std Tables, Entr B Sec, Entr C, Entr B.

- Stay: if a device is seen by the same sensor, over a certain threshold (5 minutes, just enough to distinguish stays from transits) in a given time window, we conclude the user stays at the sensor zone, doing the activity in that semantic location.
- Path: for a device that is seen by a set of sensors without large interruptions, the path is defined as its sequence of stays.
- A frequent path is a path followed by many devices/users, in relation to the whole number of paths in the dataset.
- A device/user behaviour pattern is the set of frequent paths followed by many devices/users.

III. TEMPORAL ANALYSIS

First, we performed temporal analysis of the data. The purpose of this analysis is to find behaviours related to class schedules, work days, or hours of activity. To carry out this analysis, the first step is to aggregate the data so that it is easy to analyse its temporal behaviour. We did so, using the key (time, sensor), counting the number of users (Different) that were seen by said sensor during that hour and the total number of minutes in which a user is detected. This aggregation gives us a new set of data from which several conclusions can be drawn, observable both analytically and graphically.

Fig. 2 shows the number of different users seen by any sensor at a given time. From this figure, we can draw some obvious conclusions like that the activity in the school is much greater in the working days than in the weekends, or that at night there is no activity at all, but also, other less obvious conclusions can be extracted:

- The number of people in school is greater in the mornings than in the afternoons. This can be extracted by observing any particular day, the number of different users is composed of two peaks, with a valley in the middle that coincides with the lunch time. The first of these peaks reflects the number of people in the morning (whose maximum is around 11:00 am, half the morning) which is approximately 20% higher than the second peak, which reflects the number of people in the afternoon.
- Holidays, or days without lessons: On Monday 16th, the graph’s behaviour is similar to a weekend, this is because that day was festive and there were no lessons, but the library remained open to students. Lower level activity days to 1st, 2nd, 16th (holidays), and 7, 8, 14, 15, 16, 21, 22, 28 and 29th for

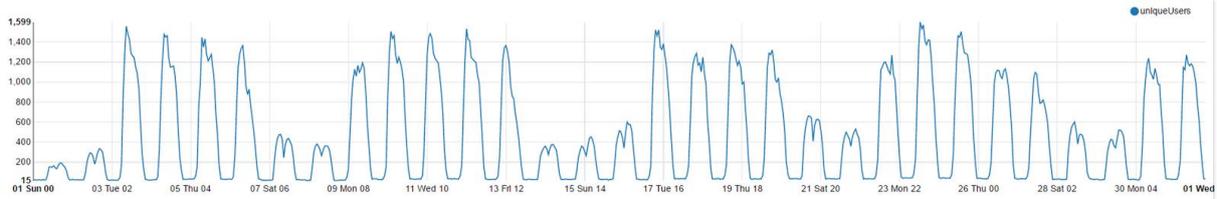


Fig. 2. Number of unique users per hour seen by any sensor

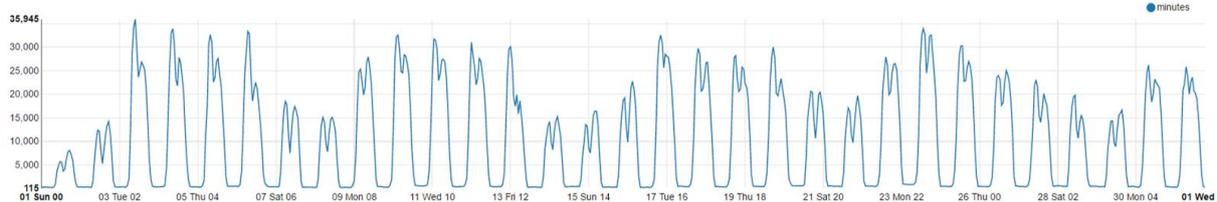


Fig. 3. Number of minutes accumulated per hour

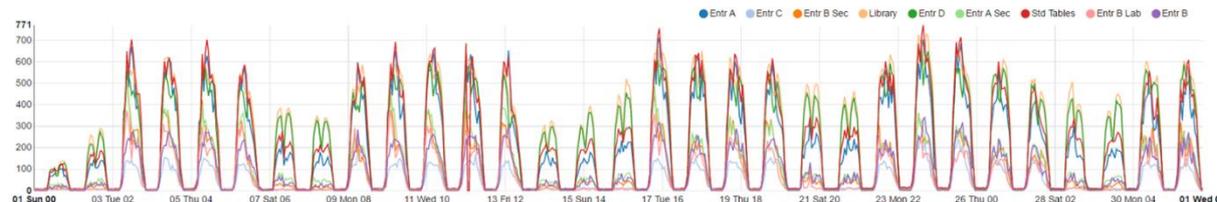


Fig. 4. Number of unique users per hour seen by any sensor

weekends. As an example of usage of library, we got 347, 336, 361, 377, 380, 354, 375, 284, 399, 349, 415 users for these days, rendering on average 361 users for a whole amount of 480 seats, which accounts for 75% of occupancy on weekends. It is also easy to see that whole activity on Mondays is lower than the rest of labour days but higher than weekends -which is due to examinations are done on Mondays.

Fig. 3 shows the number of minutes of activity accumulated in all sensors per hour. This chart allows us to distinguish between periods of transit and periods of stay. An example of this is the comparison between a weekend day and a weekday in the morning: while on weekends the number of users is much less than between daily, the relative difference between the number of accumulated minutes is not so big, this is because most of the weekend users are students in the library, who are standing for long periods of time, while on weekdays both teachers and students move through the school changing between classrooms. The same thing happens when comparing a weekday in the morning and in the afternoon, from which it can be deduced that students usually go to the library in the afternoon. Another behaviour that is observed is lunch time: at this point the graph drops to a trough, as people start to move around the school to go to eat at the cantina or outside the school.

Finally, in this analysis, Fig. 4 shows the same data as Fig. 2 but separating it by sensor. This figure, allows us to observe how users move throughout the day on an hourly basis. Observing the lines of the sensors "Entr A

Sec" and "Entr B Sec", corresponding to the sensors located near the classrooms, it is observed as the number of users grows during lesson hours. In contrast, the "Std Tables" and "Entr A" sensors are maintained at the same level during all hours of daytime activity.

IV. SPATIAL ANALYSIS

The second approach is a spatial analysis. This time, we will try to find patterns related to how people are grouped in the studied area, using the positions of the sensors as an indication of the location of the users seen. It is not intended to perform an accurate trajectory analysis, but, an analysis of buildings and areas average occupation. In this case, the data are transformed by counting the number of occurrences of the key (time, sensor, user), where, now, time is just the hour corresponding to the timestamp of the sample, regardless of the day. With this transformation, we obtain a data set in which the number of minutes that, during the month studied, a user has been seen at a certain time in a certain place can be observed; e.g. how many times the user has been seen between 10 and 11 AM by the sensor in the library.

TABLE I
UNIQUE USERS PER SENSOR

| Sensor | Unique users |
|-------------------|--------------|
| <i>Std Tables</i> | 9348 |
| <i>Entr A</i> | 8870 |
| <i>Library</i> | 8483 |
| <i>Entr D</i> | 7810 |
| <i>Entr A Sec</i> | 6770 |
| <i>Entr B</i> | 6329 |
| <i>Entr B Lab</i> | 3798 |
| <i>Entr C</i> | 2524 |

The first result allows to get an estimate of the profiles of the people detected by the sensors. Table I shows the total number of (different) users seen by each sensor throughout the month. By using the semantic location of the sensors, users can be classified by the activity that is performed in the coverage area of each sensor:

- Entr A: this is the main entrance; any kind of user can be observed here.
- Entr C: building C is composed by offices and investigation laboratories, so, most of the user seen here will be teachers or researchers.
- Entr B sec: the coverage of this sensor overlays the classrooms of building B, so users seen here will be both students and teachers.
- Entr A sec: this sensor is placed near the classrooms where the first-year courses take place, so first year students will be detected by this sensor.
- Library: the library is public access, students from this school and others will be seen here.

The second part of this analysis is based on the users' centroid analysis. This centroid is calculated as the average position of each sensor which saw the user weighted by the number of minutes seen during a defined time interval.

$$(1) \quad C_u = \frac{\sum_{i=1}^N p_i * m_{u,i}}{\sum_{i=1}^N m_{u,i}}$$

Where C_u is the centroid of a single user, N is the number of sensors, p_i is the position of the i -th sensor and $m_{u,i}$ is the number of minutes which the i -th sensor detected the user in the studied period of time. Then, these centroids are painted in a heat map over the school map to show the users concentration in each zone.



Fig. 5. Heatmaps of centroids since 8:00 to 10:00

(1) 8:00-9:00 (2) 9:00-10:00

Fig. 5 shows the heatmaps at 8 and 9 am, showing the transition between hourly intervals. This interval coincides with the start of daily activity, since the school is closed at night. The map on the left shows the users centroids between 8:00 and 9:00, “Entr A” is the sensor that registers more activity, since it is the main entrance and most of users access the area studied through this point. In the map to the right, which corresponds to the period between 9:00 and 10:00, it is clear that the number of users has been rapidly increasing. Four hot zones stand out:

- 1) “Entr A”: users are still entering the school.
- 2) “Library”: from the first hour, the library is full of students.
- 3) “Entr A Sec”: this part of building A contains most of the classrooms. The morning schedule of classes is from 9 to 13, so the users that appear around this sensor will be the students who are there.
- 4) “Entr B Sec” and “Entr B Lab”: like the above, these two sensors cover an area of classrooms, therefore, the heating of this area is also related to the time when lectures start.

Fig. 6 presents the heatmaps corresponding to the four hourly intervals which cover lunchtime, from 12:00 to 16:00. Lessons usually end at 13:00 (some end at 14:00), consequently, lunchtime is from 13:00 to 15:00 (the two central maps). The school cantina is located outside the area enclosed by the sensors. But, the path to it crosses the coverage of sensors “Entr A” and “Std Tables”, therefore, people going toward or staying at the cantina will be seen by these two sensors.



Fig. 6. Heatmaps of centroids since 12:00 to 16:00.
 (1) 12:00-13:00 (2) 13:00-14:00 (3) 14:00-15:00 (4) 15:00-16:00

Observing the transition between the maps, it is evident how the user’s centroids move towards the zone of the cantina (southwest part of the map), and between the 13:00 and the 15:00 the zone gets a higher activity than the normal one. It can also be seen how the sensor that covers the classrooms of building A, "Entr A Sec", has activity on the first two maps, on the third map, 15 to 16, it is practically turned off since there are no lectures at that time; starting at 16:00 it warms again, coinciding with the beginning of the afternoon class schedule.

V. BEHAVIOUR ANALYSIS

Finally, the third analysis performed is a behaviour analysis. On this occasion, the goal is to find behaviour patterns that describe the activities that are performed in the school and which are an indication to identify the different user profiles that can appear in the system. Final exams period begins at the end of May, so it can be expected (as can be inferred from previous analyses) that the most common behaviours will be students who stay long periods of time studying in the library.

To perform this analysis, we made a joining of sensors considering their semantic positions. The sensor "Entr D" covers the back of the library, and most of the users it detects are students in the library, being able to verify

that because users detected by this sensor, in general, are also detected by the sensor "Library"; With this premise, the data of these two sensors are joined. The second join is made between the sensors "A" and "Std Tables", due to their proximity and that both cover part of the main hall, the cantina and the stairs (transit areas).

To discover each user’s behaviour, data is partitioned by day and user. For each key (day, user) and each hour of the day along the activity hours, we extracted the sensor that has seen that user most of time. This renders the desired behaviour for that certain day and user. A total of 58311 behaviours are obtained, of which there are 18326 unique values, but analysing the most frequent it is detected that many of them are quite similar. Of the behaviours found, 42% of them (24488) only have activity in the morning, 35% (20717) only have activity in the afternoons, and 83% (48,608) are interrupted for at least one hour during mealtime.

Table II presents the 10 most frequent behaviours (the ones that are most repeated). Each row represents one frequent behaviour over the observation time (working hours), columns show, in one-hour intervals, where users that follow this behaviour are. At first sight, almost all of them can be grouped into two kinds: morning stays in the library and afternoon stays in the library.

Table II

TOP 10 MOST FREQUENT BEHAVIOURS

| Index | 9:00 | 10:00 | 11:00 | 12:00 | 13:00 | 14:00 | 15:00 | 16:00 | 17:00 | 18:00 | 19:00 | 20:00 |
|-------|---------|---------|---------|---------|---------|-------|---------|---------|---------|---------|---------|-------|
| 1 | | | | | | | | Library | Library | Library | Library | |
| 2 | Library | Library | Library | Library | Library | | | | | | | |
| 3 | | | | | | | | | Library | Library | Library | |
| 4 | Library | | | | | | | | | | | |
| 5 | | | | | | | Library | Library | Library | Library | Library | |
| 6 | | | Entr A | | | | | | | | | |
| 7 | | | | | Entr A | | | | | | | |
| 8 | Library | Library | Library | Library | | | | | | | | |
| 9 | | | | | | | | | Entr A | | | |
| 10 | | Library | Library | Library | Library | | | | | | | |

Table III

| TOP 10 MOST FREQUENT BEHAVIOURS FILTERING THE MOST COMMON SENSORS | | | | | | | | | | | | | | | | |
|---|-------------|---|-------------|---|-------------|---|-------------|---|-------------|---|-------|--|-------------|---|-------------|---|
| Index | 9:00 | | 10:00 | | 11:00 | | 12:00 | | 13:00 | | 14:00 | | 15:00 | | 16:00 | |
| 1 | Entr Sec | A | Entr Sec | A | | | | | | | | | | | | |
| 2 | Entr Sec | A | | | | | | | | | | | | | | |
| 3 | Entr Sec | B | | | | | | | | | | | | | | |
| 4 | | | | | Entr Sec | A | | | | | | | | | | |
| 5 | | | | | Entr Sec | B | | | | | | | | | | |
| 6 | | | Entr Sec | A | | | | | | | | | | | | |
| 7 | | | | | | | | | Entr Sec | A | | | | | | |
| 8 | | | Entr Sec | B | | | | | | | | | | | | |
| 9 | | | | | | | | | | | | | Entr Sec | B | Entr Sec | B |
| 10 | Entr Sec | A | Entr Sec | A | Entr Sec | A | Entr Sec | A | | | | | | | | |

The difference in activity registered by the sensors "Entr A", "Std Tables", "Library" and "Entr D", may hide the typical behaviours seen by other sensors. In order to prevent this, we filter previous result by searching only behaviours which not contains any of the most active sensors. The result of this filter is presented in Table III. The table, shows again the ten most frequent behaviours, this time with the named filter. It shows behaviours expected in a school: lecture attendance. Most subjects are imparted in lessons of two hours, starting at 9:00, 11:00, 15:00 and 17:00.

VI. CONCLUSIONS

Indoor positioning, path discovery and resource allocation have been a research area with great interest on recent years. Most of the current techniques require a precise and frequently updated position of the users. This paper proposes a simple approach to approximate to these problems using a very inexpensive infrastructure.

Semantic trajectories provide a method to get answers to some of the questions raised by these problems without the need of a precise user tracking. A simple and small network of low cost Wi-Fi sensors is enough to perform the analysis. The sensors scan every minute the MAC address of any device which has an active Wi-Fi sensor inside their coverage area. Accumulating these data for a medium period of time, one month in this analysis, results in a dataset with sufficient information to obtain some interesting conclusions.

Along our analysis' description we have proved that is possible to extract useful information about the operation of the school and about its users' behaviours, having a limited prior knowledge. Studying the temporal distribution of the number of people in the school we could distinguish between a weekday and a weekend or holiday, it is also quite easy to find out the installation working hours. We show how people from

all around the area concentrate in the cantina during lunch time. Comparing the number of unique users seen by each sensor, and taking into account their semantic position, we could estimate the library occupancy during weekends. Finally, we discovered the massive use, during all day, of the library during the month before the final exams.

In future works, we will include data from a longer observation period and increase our scope including sensors located in another schools. Increasing our dataset both temporal and spatially will allow us to discover new behaviour patterns, confirm our observations and apply our methodology to new scenarios.

REFERENCES

- [1] Czogalla, O., & Naumann, S. (2016, October). Pedestrian indoor navigation for complex public facilities. In *Indoor Positioning and Indoor Navigation (IPIN)*, 2016 International Conference on (pp. 1-8). IEEE.
- [2] Zheng, Z., Chen, Y., Chen, S., Sun, L., & Chen, D. Location-aware POI Recommendation for Indoor Space by Exploiting WiFi Logs.
- [3] A. Kurkcu and K. Ozbay, "Estimating Pedestrian Densities, Wait Times, and Flows with Wi-Fi and Bluetooth Sensors," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2644, pp. 72–82, Jan. 2017.
- [4] Y. Xu, I. D. G. Groeneveld, R. Sulzer, E. Theocharous, O. T. Willems, and M. S. Tryfona, "Determine activity based on the classified identity of users by using Wi-Fi monitoring," *Geomatics Synth. Gr. Proj. Rep.*, 2016
- [5] Kopytoff, V. Stores Sniff Out Smartphones to Follow Shoppers. [Online] MIT Technology Review, November 12, 2013.
- [6] Duque Domingo, J., Cerrada, C., Valero, E., & Cerrada, J. A. (2016). Indoor Positioning System Using Depth Maps and Wireless Networks. *Journal of Sensors*, 2016.
- [7] Zhang, X., Kim, G. B., Xia, Y., & Bae, H. Y. (2012, August). Human activity recognition with trajectory data in multi-floor indoor environment. In *International Conference on Rough Sets and Knowledge Technology* (pp. 257-266). Springer Berlin Heidelberg.
- [8] Schauer, L., Marcus, P., & Linnhoff-Popien, C. (2016, October). Towards feasible Wi-Fi based indoor tracking systems using probabilistic methods. In *Indoor Positioning and Indoor*

- Navigation (IPIN), 2016 International Conference on (pp. 1-8). IEEE.
- [9] Bowman, G., & Jaebker, K. (2013). Using commodity hardware as an affordable means to track on-site visitor flow. In *Museums and the Web*.
- [10] Parent, C., Spaccapietra, S., Renso, C., Andrienko, G., Andrienko, N., Bogorny, V., ... & Theodoridis, Y. (2013). Semantic trajectories modeling and analysis. *ACM Computing Surveys (CSUR)*, 45(4), 42.
- [11] Guo, S., Xiong, H., Zheng, X., & Zhou, Y. (2017). Activity Recognition and Semantic Description for Indoor Mobile Localization. *Sensors*, 17(3), 649.
- [12] Ying, J. J. C., Lee, W. C., Weng, T. C., & Tseng, V. S. (2011, November). Semantic trajectory mining for location prediction. In *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* (pp. 34-43). ACM.
- [13] Deb, B., & Basu, P. (2015, February). Discovering latent semantic structure in human mobility traces. In *European Conference on Wireless Sensor Networks* (pp. 84-103). Springer International Publishing.