Understanding User Daily Mobility Using Mobile and Wearable Sensing Systems

Sébastien Faye, Thomas Engel University of Luxembourg, SnT 4 rue Alphonse Weicker, L-2721 Luxembourg Email: {sebastien.faye,thomas.engel}@uni.lu

Abstract—Recent technological advances and the ever-greater developments in sensing and computing continue to provide new ways of understanding our daily mobility. Smart devices such as smartphones or smartwatches can, for instance, provide an enhanced user experience based on different sets of built-in sensors that follow every user action and identify its environment. Monitoring solutions such as these, which are becoming more and more common, allow us to assess human behavior and movement at different levels. In this article, we focus on the concept of human mobility. With the participation of 13 individuals, we carried out an experiment to discover how groups of sensors currently available in smartphones and smartwatches can help to distinguish different profiles and patterns of human mobility. We show that it is possible to use not only motion sensors but also physiological sensors and environmental data provided, for instance, by Wi-Fi. Finally, detailed study of these categories enables us to offer a way of representing the mobility of individual users, based on anonymized traces and graph theory.

I. Introduction

The rapid emergence of new technologies and the continuing expansion of networks, both fixed and mobile, promise new possibilities for understanding human behavior. Whether in smartphones, smartwatches, or specialized equipment, the miniaturization of sensors and the popularity of these devices allow both industry and science to propose valuable new models, concepts and prototypes. This network of sensors, or sensing systems, has the potential to be used in areas such as health, sports, and general user monitoring.

More specifically, issues related to human mobility and transportation systems are very well adapted to this type of system. If issues related to navigation, traffic flow optimization, fleet management or autonomous driving are hot topics, then user-centric systems and the possibilities they offer are a foundation we need to understand. User preferences and habits are indeed essential elements that significantly enhance the user experience. In this context, sensing systems such as those we explore here are ideal candidates.

In this article, we study ways in which sensors built into smartphones and smartwatches (two of the most popular devices of the moment) can be used to analyze and characterize the mobility of their users. To do this, we begin by describing in Section II a data collection that we conducted with 13 participants using the SWIPE open-source system. In Section III, we go on to study different sensor groups to investigate the advantages and disadvantages they may provide when studying user mobility: (1) motion detection,

(2) physiological and (3) environmental monitoring. The aim of this study is to use sensors and combinations of sensors not commonly used in similar work or in mobility studies, which in most cases only use accelerometers and user inputs. Finally, we open the way to discussion in Section IV and propose a new way to describe and visualize the mobility of an individual, based on conclusions made throughout the paper. This work can be used as input and background for future studies or prototypes that target the user experience.

II. METHODOLOGY

In this section, we define a methodology to obtain user data in motion. We used the SWIPE open-source platform, which is available online under an MIT license¹. As part of this paper, we make another platform available² to analyze and show a part of the dataset presented below in anonymized form.

A. Sensing System Architecture and Metrics

The sensing system we use is an Android application that collects data simultaneously on a smartwatch and a smartphone. The architecture of SWIPE consists of two parts, which are detailed in [1]. First, the smartwatch (worn on the wrist) regularly sends the data it has collected to the smartphone (carried in the pocket). The smartphone serves as a local collection point and as a gateway to access an online platform over the Internet. This platform is composed of several modules, which (1) receive data following an authentication process and (2) store, (3) analyze and (4) display it by means of a web interface. Details of the main metrics collected are listed in Table I.

TABLE I. Key metrics collected by our Sensing System

Metrics		cording & Comments pling rates	
Maximum and average acceleration $(m.s^{-2})$	30 sec.	< 1 sec.	Maximum and average value of $\alpha = \sqrt{(x^2 + y^2 + z^2)}$, where x, y and z are the acceleration along each axis of the device, excluding gravity and provided by the accelerometer (phone and watch).
Pedometer (steps)	60 sec.	~	Number of steps taken by the user, detected by the Android system as a function of the accelerometer. (phone and watch).
Heart rate	60 to		Heart rate, in beats per minute, provided
(b.p.m.)	300 sec.		by the optical heart rate sensor (watch).
Wi-Fi APs	300 sec.		Anonymized BSSIDs of Wi-Fi Access Points (phone).
Speed $(km.h^{-1})$	60 sec.		Travel speed provided by the GPS (phone).

B. Data Collection

The platform detailed above allowed us to collect data, using a smartphone (LG Nexus 5) and a smartwatch (Samsung Galaxy Gear Live), both running Android 5.1.1. Data was collected from 13 participants working at the University of Luxembourg in the same building. Each participant was systematically subjected to the same requirements: (1) Wear the devices for one day, from 10:00 to 23:59; (2) Complete an activity diary; (3) Sign an informed consent form to accept the privacy policy of the study.

Figure 1 shows the geographical distribution of our dataset over the 13 participants. These traces are largely concentrated in one area – the participant's workplace. In most cases, the remaining traces represent commuting activities or meetings in specific places.

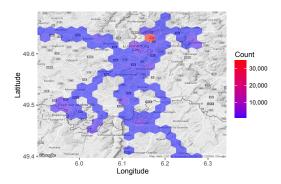


Fig. 1: Geographical distribution of our dataset

Figure 2 shows the hourly breakdown of two types of mobility over our dataset. The red bars represent physical activities, while the blue bars represent activities where the user is in a motorized vehicle. Overall, we see a fair distribution of physical activities, except in the afternoon, when, for the most part, participants are sitting in their office. Vehicular activities occur predominantly in the evening, when participants return home. Physical activity also appears to peak around 20:00, when some users participate in a sport.

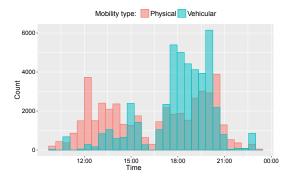


Fig. 2: Mobility distribution of our dataset

C. Energy Saving Strategy

The provision of a sensing system launched as a background service represents a potential burden on the batteries of the

devices used, which are not renowned for their longevity. As described in [1], it is therefore critical that we make every effort to save energy. By optimizing data transmission, recording frequency and the devices themselves (e.g. optimizing default running services), we find an autonomy gain of about 287% for the smartwatch (13.5h vs. 4.7h with high transmission, harvesting and recording frequencies) and on the order of 189% for the smartphone (15.7h vs. 8.3h). Figure 3 represents the energy-expenditure profile of the equipment used. The blue and the orange areas show measurement points for both device - blue for smartphones and orange for smartwatches. The trend line in the middle of each area is a local regression line representing the tendency of these recordings. Overall, we see that our system can easily work for at least 12 hours, despite all the sensors being used in conjunction with the low energy capability of the smartwatch. For its part, the smartphone obviously has better capabilities, which are, however, used quite heavily due to its role as a relay point between the watch and the Internet. These energy profiles give us good reason to believe that our system will function for longer periods of time on newer hardware with better batteries.

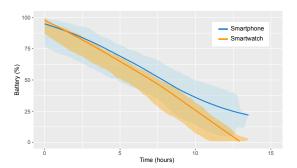


Fig. 3: Energy-expenditure profile of the smart devices

III. EXPLORING SENSOR CATEGORIES

Understanding human activity and mobility patterns is a fundamental prerequisite in providing context-aware mobility services. Traditionally, characterizing human mobility has depended solely on data available from cellular networks and GPS devices attached to the subject whose movement or activity pattern is under study. However, the proliferation of smart devices equipped with a rich set of high precision sensors has caused a paradigm shift in the provisioning of mobility and activity recognition based services. Activity recognition applications make use of these sensors to detect the physical activities the user performs such as standing idle, walking, running, or driving a car or bicycle. Several studies have been conducted in these emerging research fields; interested readers can refer to existing surveys [2], [3].

In this section, we describe three popular categories of sensors in order to provide the reader with a broad view of the relationship between mobility and sensing systems. Relying on our data set, we discuss and propose for each category original ways of analyzing user mobility. The conclusions of this section will serve as the basis for the generation of a daily mobility profile in the next section.

A. Motion Sensor Metrics

According to the literature (e.g. [4]), motion sensors are among the most used both in specific research connected with movement detection and more general studies focused on user travel patterns. These sensors, typically 3-axis accelerometers, can accurately trace the movements made by a device. For example in [5], Castignani et al. investigate how motion sensors can be used to detect risky driving events and develop a platform for monitoring driving habits. In addition to these sensors, GPS subsystems fall into this category because, apart from allowing positioning of an object or user in space at a larger scale, they provide comprehensive travel data (e.g. speed). The GPS sensor is the primary positioning sensor on smartphones. It determines a device's physical location by returning its longitude and latitude coordinates.

Figure 4(a) provides a fundamental diagram that identifies three main categories of mobility: still, physical and in-vehicle activities. At a specific time, each point of the graph conflates three values measured by the smartphone: maximum linear acceleration, average linear acceleration and GPS speed. The set of points is extracted from our data set, without distinction between users. Each point is displayed in a different color corresponding to the activity the user was performing based on data recorded at the time, validated using the activity diaries. We can see a clear trend between three types of activity. First of all, trips made by car or in a motorized vehicle (green) seem clearly independent of linear acceleration, but are closely correlated with GPS speed changes. Conversely, activities causing bodily energy expenditure further stimulate the accelerometer – to a relatively low extent when walking, becoming much stronger when running. Between the two, inactivity (blue) logically causes little response over the axes.

This explains how we can easily distinguish several classes of fundamental activities. Figure 4(b) gives us a similar graph created with data from the smartwatch. As users wear smartwatches on their wrists while traveling, we can clearly

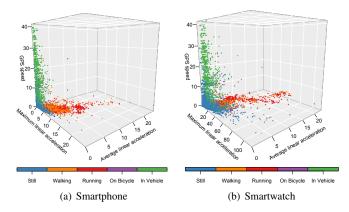


Fig. 4: Fundamental activity diagrams

see more information about sudden movements, especially those made in a vehicle.

B. Physiological Metrics

Physiological sensors can be used to capture electronic signals from the human body. Such a signal may be electrocardiographic (ECG) data which can be used to infer a user's stress and emotion level [6], [7]. This type of data can serve as an indicator for capturing driving stress experienced by users. When users are considered as groups or collectively, this information has the potential to identify critical geographical or temporal points, i.e. places that are perceived as dangerous or perceived negatively by the majority of users, as depicted in Figure 5. With this information, it would be possible to supply a mobility monitoring system that encouraged individual users or a group of users to drive on roads that avoid places they usually perceive negatively.

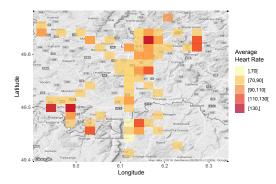


Fig. 5: Stress zones in our dataset

In terms of recording on a device, it is quite difficult to compare different absolute physiological data, because these can vary greatly from user to user. However, it is easy to imagine a system that measures relative data, such as a coefficient of variation, at regular time intervals.

As discussed in [8], in the reminder of this article, we do not consider this type of data. While physiological data seems useful for understanding stress and other feelings experienced by the user, it is too dependent on the hardware (smartwatch), consumes energy, and lacks the precision needed to provide a reliable long term solution.

C. Environmental Metrics

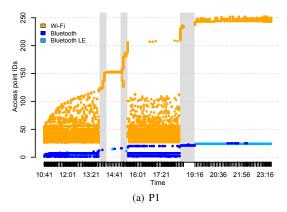
Rather than measuring the actions and reactions of users, environmental sensors monitor a user's surroundings. This category integrates sensors that measure and monitor environmental conditions such as ambient temperature, atmospheric pressure, and illumination. Such sensors include thermometers, barometers, and light sensors. Moreover, it is worth noting that Wi-Fi and Bluetooth technologies can also be used as environmental metrics able to study mobility patterns [9] and locale characteristics [10]. Several other studies have been based on the use of sensors such as microphones [11] and cameras [12] as a means of automatically recognizing particular places, or simply for identifying the context that

users find themselves in. However, analysis of audio traces requires microphones to record continuously, raising privacy issues. Recording of videos is also expensive — especially on mobile platforms where energy is limited. In this section, we propose the analysis of network traces, focusing on Wi-Fi traces and their potential for analyzing human mobility.

Figures 6 shows one of the ways that network traces allow us to understand the movements and interactions a person makes. These figures represent the evolution, over the course of one day, of interactions between Wi-Fi access points (APs) and between Bluetooth devices, using a unique numerical identifier generated from their BSSID (Y axis). Figure 6 illustrates the case of two users: P1, who does not move much when at his workplace; and P12, who moves much more. mainly between different meeting rooms and campuses. Both users drive a car. The gray areas indicate when participants are commuting. We can see that both users tend to encounter a large number of Wi-Fi APs, indicating the spatial movements of these users throughout the day. For example, for the first part of his day, user P1 remains in a certain place before moving to another place around 12:00, staying there for an hour, and then returning to the original place. User P12 appears to have much greater mobility, visiting many more places and staying there for less time. It is interesting to see that when P12 goes home around 18:00, he continuously encounters multiple Wi-Fi networks, indicating that he is moving slowly. Conversely, P1 moves faster and comes across only a few networks. Finally, both participants end their day at home (new Wi-Fi IDs).

To take things further, we wanted to offer a way of representing interactions between users and networks. To do this, we turned to graph theory. Each Wi-Fi AP that the user scans, which is usually stationary, is shown as a node. Each time the user scans two separate APs at the same time, an edge is created between the two corresponding nodes. The weight of each edge is simply the number of times the user has scanned this edge. The resulting non-directed graph provides quantitative and qualitative information on mobility and user interactions.

Using the principles described above, figures 7(a) and 7(b) represent the Wi-Fi interaction graphs of P1 and P12. We can clearly see different features. First, P1 has very distinct connected components, including a large matching with his workplace, which has many APs. A connected component models a group of nodes that are connected together, but disconnected from the rest of the network. P12 also has large groups of nodes; however, in this case there are clear links between each one, indicating slow movement between these groups of nodes or locations. Refocusing on graph theory, we can explore different traditional metrics to characterize mobility differences between these two users. Table II gives us an example of some of these metrics. The number of connected components gives an indication of the number of places visited. The "diameter" gives an indication of the size of the largest connected component and on average over all connected components, which can indicate whether or not the



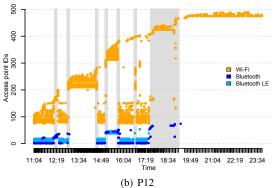


Fig. 6: Network traces represented as a temporal figure

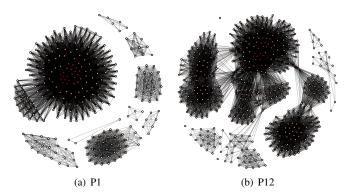


Fig. 7: Wi-Fi traces represented as a non-directed graph

TABLE II. Graph statistics (Wi-Fi)

	P1	P12
Connected components	9	7
Average diameter	1.4	1.6
Maximum connected component diameter	2	6
Average Weighted Degree	403.3	109

user is physically active inside visited places. The average weighted degree indicates how much time the user spent in each place visited, because it considers the number of times the networks were scanned.

D. Sensor fusion

To illustrate the possible correlations between certain measurements, Figure 8 shows an example of how we can rep-

resent a large number of metrics. With participant P11 as its subject, the figure displays successively (1) activities detected by a native Android algorithm, (2) the user's heart rate, (3) the linear acceleration of the smartphone and smartwatch, and (4) a variety of anonymized geographical information. This figure allows us to easily understand the different relationships between sensors and visualizes the user's main activities throughout the day. It is interesting, for instance, to note the relationship between linear acceleration and the detected activities. For example, around 10:00, linear acceleration is detected by the smartwatch alone, suggesting that the user is stationary but moving his arm – probably at his desk. Between 19:02 and 20:12, the user is clearly detected as moving (GPS), but with a low linear acceleration over the two devices. This tends to validate the detected activity, which is being in a motorized vehicle. This is in contrast with the activity between 20:12 and 21:22, where the high linear acceleration suggests that the user is running.

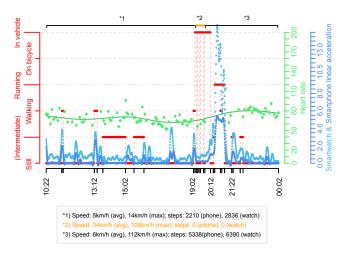


Fig. 8: Example of multiple sensor data representation (P11)

IV. EXTRACTING MOBILITY PROFILES AND PREFERENCES

As we have seen, several metrics can be taken into consideration, and the relationships between them can be used to compute aggregated values such as activities or locations. In this section, we use the conclusions arrived at in Section III in order to see how it is possible to create a representative and visual daily mobility profile for a user.

As this profile is intended to be both simple and energy-efficient to implement, we should avoid using too many different devices or sensors. Thus, we make the following choices: provide a modular solution, which will, by default, include metrics provided exclusively by the smartphone; avoid environmental data collected by microphones or cameras because of computational and energy constraints; and use anonymized metrics to avoid compromising the privacy of the user. Taking these choices into account, the graph theory aspects studied above seem to be ideal candidates because they are easy to retrieve, inexpensive to obtain in terms of energy (i.e. recording frequency does not need to be high) and they

are easily generalizable to several time scales (e.g. hour, day, week). Thus, the profile we generate uses only a smartphone, recording two groups of metrics constantly and at equal intervals: (1) anonymized (i.e. hashed) Wi-Fi AP BSSIDs encountered by the smartphone, in order to build for each user a graph G similar to the one presented in section III-C; and (2) activities performed by the user, only as labels (i.e. post-computed data). Finally, a relationship is created between those two groups by associating a Boolean with each node to indicate whether or not the user was mostly moving (either physically or through some form of transportation). The concrete implementation of the application is detailed in [13].

This profile is build on top of **four aspects** that we consider essential when profiling human mobility:

 \square **Number of visited locations**. We extract this metric from the sub-graph G_{still} , which is composed only of nodes where the user was not in motion. To represent this metric, we suggest using the number of connected components of G_{still} divided by a unit time. Looking at similarities with visited locations computed using GPS data, we found a Pearson's correlation coefficient of 0.93, which confirms the adequacy and accuracy of this value.

☐ **Urban index**. The second aspect we want to consider is the paths between these different places. For this, we consider the complement graph of G_{still} , namely $G_{mobility}$, which only consists of nodes with moving activities. Based on this graph, we extract what we call an urban index, representing whether the environment in which the user is moving is an urban area. We represent it by using the number of nodes in $G_{mobility}$ per activity unit time: the more a user moves around in an urban environment, the more he will tend to come across Wi-Fi APs. In order to validate this index, we computed a similar metric using GPS locations and OpenStreetMap data, defining that a certain GPS location is in a urban area when there is at least one residential road within 420 meters (tested empirically). This GPS index is then equal to the proportion of urban areas. Looking at similarities with our urban index, we found a correlation coefficient of 0.74, which is not perfect, but good enough to validate our metric, which has a slightly different meaning.

 \Box **Spatial exploration**. Looking at the complete graph G (i.e. $G_{still}+G_{mobility}$), we compute an index showing whether or not the user has a tendency to move inside the most important place he visited. This metric has a relationship with the number of articulation points of the maximum connected component. The maximum connected component is the connected component that contains the largest number of nodes, while an articulation point is a node whose removal disconnects the component it belongs to. In order to validate this index, we compute the average duration of the walks of each user when he is at his workplace (i.e. the maximum component). The longer the user moves in a place, the more he will tend to "explore" this place. Comparing these values to our index, we found a correlation coefficient of 0.76, which also confirms, to a certain extent, the validity of this index.

☐ **Activities**. Finally, separate from the graph theory prin-

ciples, we compute two indexes for each mobility activity: physical and in-vehicle activity. These scores simply reflect the proportion of time the user was performing a physical activity or was in a vehicle. They are simply computed by dividing the count of each activity over the total count of all activities. The accuracy is high and was verified using the activity diaries provided by the users and the timestamped data.

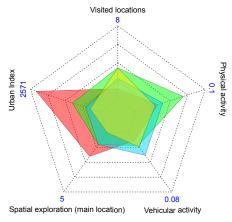


Fig. 9: Mobility profile of P1 (yellow), P7 (red), P11 (blue), and P12 (green)

Figure 9 shows four examples of profiles in the form of a spider graph based on data collected on four participants having different characteristics in our dataset. The minimum and maximum axes for each characteristic are determined according to our dataset. Participants P1 (yellow), P11 (blue) and P12 (green) drive a car and live in cities which are far from their work, while participant P7 (red) lives near to his workplace and gets around by bike and by public transport, as reflected by his high urban index (i.e. a large number of Wi-Fi APs scanned when moving). As introduced in Section III-C, participant P12 has a sporty profile, as reflected by his high physical activity score. Conversely, participant P1 does not move much when at his workplace, as reflected by his relatively low spatial exploration index. Note that participant P1 is an intermediate case, having results situated in the middle of our experiments. Finally, participant P11 travels more than other participants, as reflected by his vehicular activity score.

V. CONCLUSION

In this article, we have explored different ways of studying user mobility, proposing creative ways in which the study can be achieved. We used an open-source platform and an anonymous dataset collected over 13 participants. In particular, we used network and activity data acquired from smartphones and smartwatches. We show that the use of anonymized network data in parallel with graph theory has a good potential for describing and characterizing a user's mobility. Physiological and motion data is also useful in characterizing specific aspects of user mobility behavior. Using these findings has allowed us to propose a way of describing user mobility by creating a profile in the form of a spider graph. This profile offers

a simple and inexpensive way to analyze five characteristics showing user preferences and mobility choices.

In future work, we intend to take advantage of these conclusions, which provide clear inputs to support intelligent mobility systems such as navigation services. It is easy to imagine this kind of service taking into account different types of profiles (e.g. automatic selection of mode of transportation, favorite route and places). While the profile generation method suggested in this paper remains an illustrative example of the potential for sensor systems, we also plan to formally validate the relationship between graph theory and mobility patterns, in addition to introduce new graph theory aspects, such as dynamic graphs and time-dependent components.

ACKNOWLEDGMENT

The authors would like to thank the National Research Fund of Luxembourg (FNR) for providing financial support through the CORE 2013 MAMBA project (C13/IS/5825301).

REFERENCES

- S. Faye, R. Frank, and T. Engel, "Adaptive activity and context recognition using multimodal sensors in smart devices," in *The 7th International Conference on Mobile Computing, Applications and Services (MobiCASE'15)*, Berlin, Germany, Oct. 2015.
- [2] O. D. Incel, M. Kose, and C. Ersoy, "A review and taxonomy of activity recognition on mobile phones," *BioNanoScience*, vol. 3, no. 2, pp. 145– 171, 2013.
- [3] Ö. Yürür, C. H. Liu, Z. Sheng et al., "Context-awareness for mobile sensing: a survey and future directions," *IEEE Communications Surveys* & *Tutorials*, vol. 18, no. 1, pp. 68–93, 2014.
- [4] P. H. P. Chao Xu and P. Mohapatra, "Finger-writing with smartwatch: A case for finger and hand gesture recognition using smartwatch," in HotMobile 2015. ACM. ACM, 2015.
- [5] G. Castignani, T. Derrmann, R. Frank, and T. Engel, "Driver behavior profiling using smartphones: A low-cost platform for driver monitoring," *IEEE Intelligent Transportation Systems Magazine*, vol. 7, no. 1, pp. 91– 102, 2015.
- [6] G. Rigas, Y. Goletsis, and D. I. Fotiadis, "Real-time driver's stress event detection," *IEEE Transactions on intelligent transportation systems*, vol. 13, no. 1, pp. 221–234, 2012.
- [7] K. Soman, V. Alex, and C. Srinivas, "Analysis of physiological signals in response to stress using ecg and respiratory signals of automobile drivers," in Automation, Computing, Communication, Control and Compressed Sensing (iMac4s), 2013 International Multi-Conference on. IEEE, 2013, pp. 574–579.
- [8] S. Faye and R. Frank, "Demo: Using wearables to learn from human dynamics," in *Proceedings of the 13th Annual International Conference* on Mobile Systems, Applications, and Services. ACM, 2015, pp. 445– 445.
- [9] T. Ellersiek, G. Andrienko, N. Andrienko et al., "Using bluetooth to track mobility patterns: depicting its potential based on various case studies," in Proceedings of the Fifth ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness. ACM, 2013, pp. 1–7.
- [10] H. Liu, J. Yang, S. Sidhom et al., "Accurate wifi based localization for smartphones using peer assistance," *IEEE Transactions on Mobile Computing*, vol. 13, no. 10, pp. 2199–2214, 2014.
- [11] L. Ma, D. Smith, and B. Milner, "Environmental noise classification for context-aware applications," in *Database and Expert Systems Applica*tions. Springer, 2003, pp. 360–370.
- [12] L. Porzi, S. Messelodi, C. M. Modena, and E. Ricci, "A smart watch-based gesture recognition system for assisting people with visual impairments," in *Proceedings of the 3rd ACM international workshop on Interactive multimedia on mobile & portable devices*. ACM, 2013, pp. 19–24.
- [13] S. Faye, I. Tahirou, and T. Engel, "Demo abstract: Human mobility profiling using privacy-friendly wi-fi and activity traces," in *The 14th* ACM Conference on Embedded Networked Sensor Systems (SenSys 2016), Stanford, CA, USA, Nov. 2016.