IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

SPECIAL SECTION ON FUTURE GENERATION SMART CITIES RESEARCH: SERVICES, APPLICATIONS, CASE STUDIES AND POLICYMAKING CONSIDERATIONS FOR WELL-BEING [PART II]

Received June 16, 2020, accepted July 9, 2020, date of publication July 20, 2020, date of current version August 4, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3010609

A Privacy-Aware Crowd Management System for Smart Cities and Smart Buildings

JUAN RAMÓN SANTANA, LUIS SÁNCHEZ, PABLO SOTRES, JORGE LANZA, TOMÁS LLORENTE, AND LUIS MUÑOZ, (Senior Member, IEEE)
Network Planning and Mobile Communications Laboratory, Universidad de Cantabria, 39005 Santander, Spain

Corresponding author: Juan Ramón Santana (jrsantana@tlmat.unican.es)

This work was supported in part by the Spanish Government (MINECO) by means of the Project Future Internet Enabled Resilient CitiEs (FIERCE) under Grant RTI2018-093475-A-I00, and in part by the European Union's Horizon 2020 Programme through the European project Federated CPS Digital Innovation Hubs for the Smart Anything Everywhere Initiative (FED4SAE) under Grant 761708.

ABSTRACT Cities are growing at a dizzying pace and they require improved methods to manage crowded areas. Crowd management stands for the decisions and actions taken to supervise and control densely populated spaces and it involves multiple challenges, from recognition and assessment to application of actions tailored to the current situation. To that end, Wi-Fi-based monitoring systems have emerged as a cost-effective solution for the former one. The key challenge that they impose is the requirement to handle large datasets and provide results in near real-time basis. However, traditional big data and event processing approaches have important shortcomings while dealing with crowd management information. In this paper, we describe a novel system architecture for real-time crowd recognition for smart cities and smart buildings that can be easily replicated. The described system proposes a privacy-aware platform that enables the application of artificial intelligence mechanisms to assess crowds' behavior in buildings employing sensed Wi-Fi traces. Furthermore, the present paper shows the implementation of the system in two buildings, an airport and a market, as well as the results of applying a set of classification algorithms to provide crowd management information.

INDEX TERMS Smart city, Internet of Things, crowd management, artificial intelligence, positioning.

I. INTRODUCTION

Nowadays, modern technologies are enabling and widening the connectivity possibilities all over the world [1]. Furthermore, during the forthcoming years, we will be witnesses of a new era in which a vast majority of the traffic will be produced by Machine-to-Machine communications. According to Cisco, IoT devices will represent half of the Internet connections worldwide, with a total of 14.6 billion IoT connections by 2022 [2]. Certainly, such massive deployment of IoT devices will have a positive impact in multiple domains, being the Smart City paradigm one of the application areas that has attracted more attention. At present, IoT-based technologies are already transforming traditional public services [3] such as urban planning, lighting or waste management, just to name a few.

In this regard, one of the services that is prone to be benefited from the disruption of IoT technologies is crowd management [4]. Cities are experiencing an enormous increase of population. According to the UN's Department of Economic and Social Affairs, 68% of the population will be living in cities by 2050, compared to the 55% of today [5]. World's urbanization rate is arising multiple problems, such as sustainable infrastructure developments or crowd control. Either under normal circumstances or at large events (e.g. concerts, sports or emergencies), control of large populations in urban areas require a deep knowledge of crowds' situation and behavior. Besides, in the light of recent events that have rocked the world due to the COVID-19 disease outbreak, crowd management has become also fundamental in critical infrastructures (e.g. airports) or leisure areas (e.g. parks or markets). This paper presents a solution that is based on the belief that IoT technologies can effectively contribute to the management of crowds in cities.

There are three key contributions presented in this paper, as summarised as follows. Firstly, we present a system architecture that is the baseline for implementing novel crowd management solutions based on the analysis of captured

The associate editor coordinating the review of this manuscript and approving it for publication was Miguel Jesus Torres Ruiz.

Wi-Fi frames originated from people's smartphones. This architecture is meant to be easily replicated, providing a common and extendable framework for future implementations of Wi-Fi-based solutions for crowd management. Secondly, a particularly important feature that is described in the paper is how the proposed solution addresses the existing European Union privacy regulation regarding MAC address usage, which has to be legally binding. In this sense, the paper describes an effective solution to properly anonymise this datum for real-time crowd management. Finally, the paper presents two real-life deployment scenarios in Santander, including an airport and a city market. In this sense, the paper not only describes the implementation, but also the results of applying several statistical classification models for crowd assessment in critical infrastructures to validate the implementation of the crowd management system that has been carried out.

The remaining of the paper is structured as follows. Section II reviews the existing literature of crowd management solutions, with special emphasis on Wi-Fi-based solutions, focusing on the key features of the novel solution for a system architecture described in the paper. Section III outlines SmartSantander, the Smart City framework where the work has been carried out. Section IV presents the proposed system architecture for supporting crowd management services as well as the key design features that have been considered. Section V describes the real-world deployment carried out in two critical infrastructures of Santander, namely a public market and the airport, and the software implementation details. Section VI introduces the statistical models that have been developed and tested in the aforementioned deployments. Finally, the paper presents in section VII the main conclusions derived from our work.

II. RELATED WORK

Most of the deployed systems to monitor people's flow are based in a reduced set of technologies, including: video cameras, radio frequency, information from smartphone sensors and crowdsensing, or radio technologies, such as Bluetooth or Wi-Fi. All these technologies offer different advantages based on their characteristics.

Crowd management systems based on video cameras present important advantages compared with Wi-Fi-based solutions, such as the possibility of reusing existing infrastructure (e.g. traffic cameras); the possibility of increase their accuracy based on new developments in image recognition [6]; or the near real time feedback provided. However, this technology is highly dependent on visual aspects (e.g. weather conditions, environment changes, etc.), that could incur in extra costs associated to its deployment depending on the area to be monitored. Besides, they have to deal with the privacy concerns [7] that are inherent to the use of video cameras, which is extremely restricted in many places, even if the data is properly anonymised.

On the other hand, Radio Frequency systems provide highly accurate results in the short range and their use does not introduce any major privacy issue. Furthermore, they do not depend on the users as Wi-Fi-based system do (i.e users need to carry a smartphone to be detected). Their operation is based on the signal reflections produced by human bodies when a signal is directed against them [8]. Despite their accuracy, they do not perform well in crowded areas, which is the main focus in crowd management in cities. Moreover, they are not able to distinguish between individuals, which makes almost impossible to assess flows of people in public buildings or cities.

The usage of technologies included in modern smartphones is also common in the literature to perform crowd management and monitoring. Similarly to the system presented in this paper, these solutions require a smartphone to detect individuals. Smartphones, which almost everybody carry with himself or herself nowadays, are equipped with several wireless technologies such as mobile broadband (e.g. 4G, LTE, etc.), Bluetooth or Wi-Fi, as well as GPS. Crowd management and positioning solutions based on these wireless technologies can be divided in two main categories: passive and active systems. The main difference between them lies in the need of a specific action from the users to trigger the tracking in the latter ones.

Among the active systems, we can highlight those based on the GPS information provided by the mobile phones [9]. However, these techniques require the installation of specific apps by the user, which might deal with an increased battery consumption. Additionally, GPS location data can be inaccurate indoors, limiting the crowd management to specific use cases. On the contrary, this solution provides a accurate data outdoors and does not depends on external infrastructure to work once the user has the application installed in the device.

Similarly, Pedestrian Dead Reckoning methods [10] [11], which uses internal smartphone sensors, such as the accelerometer or the gyroscope, requires the installation of an application on the smartphone. This method consists in inferring the current location of users based on their previous positions and the smartphone sensor data. However, this estimation method falls into cumulative tracking errors that reduces the precision considerably. Within this group, it is worth mentioning the crowdsensing-based system, which aggregates information from multiple smartphones to perform crowd management. This method can provide information about the crowd behaviour [12], [13], such as the estimation of queue times to access buildings or public areas (e.g. waiting time in certain restaurants). Similarly to the GPS-based solution, it does not require of external infrastructure to work.

Employing information from the smartphone's nearby base stations to which they are attached is one of the first examples of passive technology for crowd management in cities [14]. This method provides coverage in most of the cities. However, such information can only be gathered by the operator of the network, and the main drawback is the error estimation of several tens of meters at best [15]. Future radio

communication technologies promise a higher accuracy when using this method with the introduction of the 5G [16] networks.

Finally, crowd positioning and assessment solutions exploiting local and personal wireless networks such as Wi-Fi and Bluetooth have been thoroughly explored during the past years. Those using Bluetooth typically utilizes the beacons sent by the device to locate it within a specific area. Although the accuracy can be of around 1 meter, they requires the deployment of a dense infrastructure due to Bluetooth's short physical range. On the other hand, Wi-Fi-based systems have been one of the most studied methods to locate people in the latest years [17], and have a larger range compared to bluetooth-based-systems. There are several approaches to locate through Wi-Fi:

- Received Signal Strength Indicator (RSSI) methods. They determine the distance between the user device and the access points based on the measured RSSI.
- Fingerprint methods. Although they use the RSSI as well, the approach is slightly different to the previous method. In this case, the location is obtained based on the surrounding access points detected by a smartphone for a specific location in previous measurements.
- Angle-of-Arrival methods. They are based on the MIMO technology included in modern Wi-Fi interfaces that use an array of antennas. They measure the phase shift between the antennas of the array to calculate the position of the user device based on triangulation.
- Time of Flight based methods. They consist in the calculation of the distance between the receiver and the transmitter from the signal propagation time.

Despite the technological advances and the large scientific production on the positioning research topic, the existing literature on crowd management systems, tested under real conditions outside from the laboratory, is limited. In [18], authors presents a system to let people know where they are located, but it is only based on a prefilled database with the closest access points, thus not considering external monitoring of crowds as we do for crowd management. References [19] and [20] analyse the number of devices connected to different access points within the different areas of a campus, but the experiments are limited to offline analysis of extracted datasets, compared to the system presented in this paper, which is meant to deliver information about crowds in near real time. Also offline, in [21] data from surrounding Bluetooth and Wi-Fi devices is gathered from mobile phones and compared with a prefilled database. The work presented in [22] is interesting as it follows a similar approach and present an implementation architecture, including a real-life deployment, but it employs specific devices (tags) to locate people, thus limited to people with a tag. The solution presented in [23] is interesting as it uses heterogeneous sources of information, including smartphone sensors, but does not delve into the architecture for a long-term system usage. One of the most similar approaches to the system and implementation described within this paper is presented in [24]. However, it does not describe in detail the architecture followed to gather data in near-real-time. Moreover, it does not address the privacy implications of using MAC addresses neither. Likewise, [25] proposes a solution based on Kalman filters for occupancy counting using Wi-Fi probe requests, but it also lacks of information about the system architecture and the privacy considerations for a near-real-time system.

Privacy has been always an important concern while managing crowds. In this regard, literature has addressed this by trying different methods to count people in different areas. Reference [26] uses cheap environmental sensors to estimate the size of the crowds in commercial centers, but the resolution is limited to distinguish between normal and overcrowded situations. Therefore, it does not provide the advantages to test new algorithms to locate and analyse crowds in specific areas.

In general, solutions for crowd assessment and management based on Wi-Fi traces described in the existing literature do not include the functional and network architecture to deploy them in a real environment. Usually, articles analyze different positioning and classification algorithms through an offline process of a set of files that store the Wi-Fi frames captured through a measurement campaign. These solutions do not address important aspects such as the need for managing large amounts of data, providing near-real-time results, scaling the solution to large and heterogeneous areas, or satisfying the demanding privacy requirements imposed by the recent updates of the European regulation in terms of personal information.

III. SmartSantander: A SMART CITY FRAMEWORK

The SmartSantander project [27] envisioned the creation of a large-scale experimental facility for research and experimentation on architectures, key-enabling technologies, services and applications for the IoT domain applied to the Smart City branch. This objective resulted in the deployment of more than 12000 IoT sensors in the city of Santander, comprising the following domains:

- Environmental monitoring: the prevalent part of the testbed, as most of the IoT sensors belong to this domain. It includes the fixed sensors, deployed either in the lamp-posts or in the façades, measuring sensor data such as temperature, noise or luminosity. It is worth mentioning that one part of such fixed deployment can be found in the parks and gardens, so as to help with irrigation tasks by measuring air humidity or soil moisture tension, among others. In addition to the fixed IoT sensors, a set of mobile IoT sensors were also installed in the top of public transportation vehicles, such as buses. These sensors had the possibility of extend the area of measurements, and provided data from gases such as CO, NO2 or O3.

- Parking Monitoring: with more than 300 IoT devices deployed in the city center, these sensors, buried under the asphalt in the parking spot areas, measure the electromagnetic field to detect whether a car is parked on top of them or not, sending this information to the SmartSantander testbed.
- Traffic management: SmartSantander counts with installation of IoT sensors dedicated to the monitorization of traffic in the most important entrances and way outs of the city. These IoT sensors provide information about the number of vehicles per minute or the occupancy in the lanes. Additionally, existing deployments in Santander were also integrated into the testbed, providing traffic information in several roads within the city.

Apart from the physical deployments explained above, the SmartSantander project introduced the concept of ''Citizen Sensors'', which transformed citizens into sensors with several applications developed for smartphones. The ''Participatory Sensing'' app provided a mean for the citizens to send information about the city status, including cultural events or problems (e.g. broken bin, dirtiness, etc.). Additionally, it also allowed users to send information about their smartphones (e.g. luminosity sensor, accelerometer, etc.) to contribute to the SmartSantander testbed datasets. A second app was published under the name of ''SmartSantanderRA''. This app provides information about the city to their users (e.g. available parking, shops and monument information, bus information, etc.) using Augmented Reality features as the main interface. At the same time it was released, a large number of NFC tags were deployed, that could be read using the smartphone application to gather information about the specific point where the tag was placed.

The SmartSantander ambition iterated over the time, including new experimental validation in areas that were not foreseen in the inception of the project, such as the socio-economical acceptance of new solutions based on IoT technology [28]. It is worth mentioning the use of SmartSantander as testbed to test and deploy new prototypes. One interesting example of such use is the European project LEXNET [29], which one of its main goal was to study the impact in the city of various RF transmitter and the electromagnetic field radiation. Such objective required the development of low-cost RF sensors, which were tested and studied in the SmartSantander testbed.

Nowadays, the SmartSantander testbed is also part of wider movements to federate these types of facilities among Europe. Here we can highlight the FIESTA-IoT project [30], which follows an Experiment as a Service [31] approach to access different heterogeneous testbeds from a common single entry point, using cutting edge semantic-based access tools. On the other hand, FESTIVAL [32], an EU-Japan collaborative project, also provides a single platform to access heterogeneous testbeds, including not only IoT-based testbeds but also Living Labs, Open Data platforms and virtual machine providers.

IV. CROWD MANAGEMENT SYSTEM ARCHITECTURE

A. KEY DESIGN CONSIDERATIONS

Before presenting the functional architecture for the crowd management and assessment solution, it is important to briefly summarize which have been the key considerations that have driven the design and implementation decisions.

1) REDUCED TIME FOOTPRINT

Crowd management applications require quick adaptation to the conditions of the monitored area. Thus, the technology used has to be able to provide fast assessment of the crowd composition and behavior. Not only the network architecture has to be carefully considered in order not to introduce undesired delays or being affected by network congestion, but also the techniques used for the crowd assessment have to have short convergence time in order to come up with, as accurate as possible, estimations in the shortest time possible. In this sense, in the trade-off between accuracy and speed, the former might have a lower weight.

2) DIMENSIONALITY REDUCTION

Wi-Fi-based solutions make use of the IEEE 802.11 frames that they can sniff from the air to calculate the position, number, direction, etc. of the devices that have generated them. When the decision mechanisms are applied offline, the size of the analyzed dataset is not a major issue. However, if the solution is to be applied to crowd management applications, offline analysis is out of scope. Online estimation of the crowd composition requires that, from the continuous stream of information that can be sniffed at a public area or critical building within a Smart City, only the key features extracted from the Wi-Fi frames that are captured on-site are transferred to the AI algorithms making the assessment.

3) PRIVACY PRESERVATION

Crowd management is surely a key functionality of utmost importance and relevance for addressing vital challenges related with safety and emergency handling. However, increasing security does not imply a blank check. Other fundamental rights of the people have to be taken into account. The right for keeping control of their personal data have been extraordinarily protected, at least in Europe, with the General Data Protection Regulation (GDPR) that applies to all the European Union countries since last year. In this sense, the hardware addresses of their cellphones are, somehow, digital identities of their owners. Thus, it is mandatory to prevent privacy leakages. This obligation not only implies that actual MAC addresses are never known outside the local deployment but also that robust procedures are used to completely anonymize the data streams coming from the Wi-Fi frames sensors.

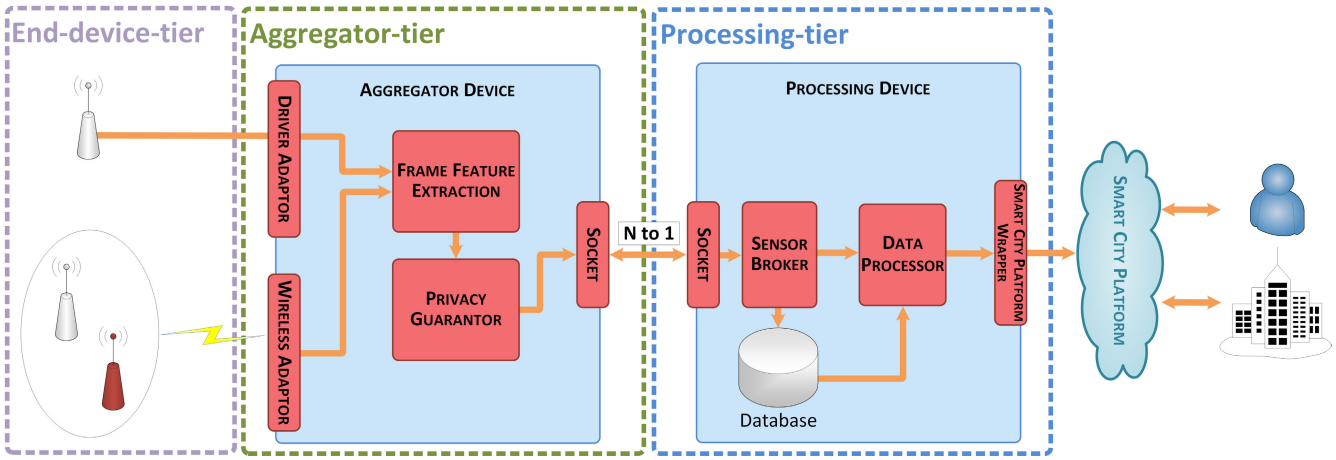

FIGURE 1. Architecture for crowd assessment and management Wi-Fi-based systems.

B. IoT DISTRIBUTED HIERARCHICAL ARCHITECTURE

The architecture shown in Figure 1 is designed to address the aforementioned considerations while enabling the provision of a high quality crowd assessment and management service based on captured Wi-Fi frames at indoor and outdoor scenarios of a Smart City deployment.

The architecture is based on a three-layered approach. Firstly, the end-device-tier, which is solely intended for raw data capture. Secondly, the Aggregator-tier, which encompasses the components in charge of data anonymization and data dimensionality reduction. Finally, the Processing-tier provides data processing capabilities to apply positioning or crowd assessment techniques on top of the previously filtered data.

The solution proposed have been designed in a deployment-agnostic manner. However, the preferred deployment model combines both fog and cloud computing. In this sense, most of the functional blocks identified can be run at edge servers or even IoT devices and only, if necessary, outsource complex AI analytics about crowd estimation to the cloud. Indeed, in the real-world deployment described in Section V, a local cloud infrastructure is used but all the AI-based prediction algorithms used could have been run at edge servers. More details about the top two tiers follow.

1) AGGREGATOR-TIER

At the Aggregator layer, raw data gathered by Wi-Fi sensors is filtered and processed after it has been anonymized for preserving people's privacy. Extracted features resulting from this processing are sent using lightweight binary frames over a raw TCP socket. While the devices running the functional components included in this tier must be continuously processing the Wi-Fi frames in order to filter and extract the required features for the crowd estimation algorithm, they do not have to have large processing capabilities as all the processes have been defined to impose the lowest computational footprint possible. This design favors the deployment of this tier leveraging fog-computing capabilities. Hence, data dimensionality reduction and low latency considerations are fulfilled since only key features are forwarded after fast and local, thus not extra-delayed, processing.

2) PROCESSING-TIER

At the top of the architecture, all the pre-processed information coming from the Aggregator-tier is finally fed into the component implementing the algorithms for crowd estimation. Additionally, the data stream is conveniently stored as it can be useful for training of deep learning techniques. Finally, results from the Machine Learning (ML) and AI techniques, this is the crowd estimation information, are formatted into the appropriate information model in order to provide it to the corresponding Smart City platform so that smart services and applications can subsequently use it.

There are 3 main components at the Processing-tier. Firstly, the Sensor Broker, which has to merge the different frames from the devices to generate a single piece of information. The frames from the same cellphone are kept in a buffer, which is released upon the reception of a frame from this smartphone with a different sequence number. The information gathered is stored in a database for persistence. Secondly, the Data Processor component hosts the ML and AI techniques that are executed over the received frames. To this end, this component combines the historical data stored in the database to create or update the model used and apply it to the frames received from the Sensor Broker in real-time. Finally, the model output is formatted and forwarded to the Smart City platform that finally exports the crowd assessment towards the corresponding crowd management application or service.

C. PRIVACY-AWARE PLATFORM APPROACH

Privacy is one of the main aspects that have to be considered when implementing Wi-Fi-based crowd estimation systems. Therefore, the architecture described in this paper takes into account the privacy considerations presented in [33]. The

privacy considerations taken are based on the recent General Data Protection Regulation (GDPR) [34] that has been set for mandatory observation, and can be considered as one of the most restrictive in the world. Assuming that requesting active permission to users so as to store and process the MAC address is not feasible, it is mandatory to remove any privacy-sensitive data from the equation, ensuring no personal data is processed and stored.

In this sense, there are two main options to remove privacy-sensitive data, such as the MAC address, either using anonymization or pseudonomyzation. There is a subtle difference between both [35]. Whilst an Anonymisation technique is a procedure of data sanitization consisting on making it no longer identifiable after this process, in pseudonymisation the true identity is deterministically replaced with an alternative identity. Hence, an adversary could eventually access to the original datum through inference attacks, or even obtain other information related to the users, such as their common routes, if there is a data breach.

The followed approach to guarantee privacy is focused on the anonymization of MAC addresses, as they are considered personal data under the GDPR. The anonymization procedure is based on the Spanish Personal Data Protection Laws and the Spanish Law Protection Office recommendations for data anonymization [36]. This document recommends the use of Hash-based Message Authentication Codes (HMAC), a cryptographic hash function, using randomly generated keys to anonymize MAC addresses.

As mentioned before, the architecture depicted in Figure 1 has been designed to leverage the fog computing concept. Aggregator-tier devices implement the privacy safeguard mechanisms before key frames' features are sent to the Processing-tier to apply the crowd estimation algorithms. The edge devices running the Aggregator-tier must be able to perform real-time filtering and cryptographic functions but these are not so computational-demanding tasks that they cannot be properly handled by regular embedded PCs.

Firstly, the Frame Feature Extraction module discards any other Wi-Fi frame but the Probe Requests. Afterwards, it extracts the key features from these frames, namely the source MAC address, RSSI, Timestamp and Sequence Number. Finally, the Privacy Guarantor component anonymize the MAC address before the information is sent to the Processing-tier server. Anonymization is performed applying the HMAC SHA256 hashing function along with a 12-bytes random key over the MAC address. The use of HMAC avoid the risks of brute force attacks, as short values are easily reversible due to the lack of input entropy. Moreover, it can be concatenated with the Sequence Number (i.e. Seq. Num. || MAC address). Concatenating the Sequence Number improves the freshness of the anonymization.

Several considerations have to be highlighted, so as to consider the case of a semi-honest adversary that might want to obtain personal information from the system, such as daily-based routes of the same individual or the MAC address using brute-force attacks. Firstly, the key used in the anonymization process must have a limited session period to be non-reversible. It must be shared among all the Aggregator-tier devices within the same deployment, so the processing layer can combine the information related to the same captured frame. Keys are randomly generated at the Processing-tier and shared among the Aggregator-tier devices, and they must be diligently destroyed afterwards. Furthermore, this key cannot be stored neither in the Processing-tier, nor in the Aggregator-tier, and the communication between Aggregator-tier devices and Processing-tier must be also kept private using secure sockets (e.g. SSL/TLS or Elliptic Curve Cryptography) and/or HTTPS protected web services. On the other hand, the generated HMAC is truncated to send only the 8 most significant bytes to ensure that a non-reversible hash is stored, even if the key is known. It is true that truncated values are more likely to be not unique, but in this case the possibility of collisions can be neglected.

The privacy-by-design considerations presented in this section have an impact on the platform possibilities. In this regard, the platform does not provide tracking information of the same individual over time. This is due to the diligent policy of key destruction per session, producing a different non-reversible identifier each time. Similarly, privacy measures taken by modern smartphone models applying MAC randomization methods will have no impact in the data provided by the system. As it is meant to locate or count users within specific areas, without keeping track of them throughout the time, an anonymized device will be located in the same area/position even if the MAC changes.

Furthermore, these considerations can also have an impact on the system scalability. The main issue regarding to system scalability can be the high concentration of people in the areas where the system is deployed. In this sense, probe requests sent by users devices can overload the systems, and the embedded PCs supporting the Aggregator-tier might not be able to handle it. This is tackled by limiting the number of probe requests processed per devices and unit of time. In this scenario, the system will be able to handle thousands of devices, just limited in this case by the database and the connection link rate between the Aggregator-tier and the Processing-tier.

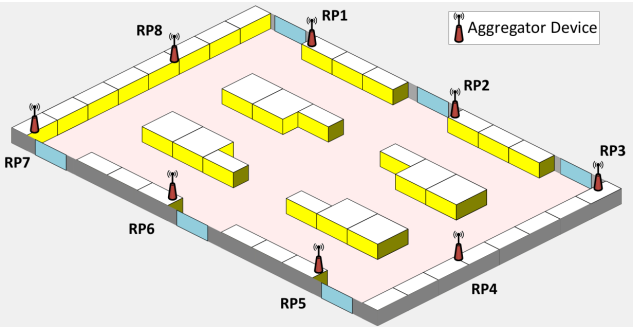

FIGURE 2. Mercado del Este deployment.

V. SYSTEM IMPLEMENTATION AND REAL-LIFE DEPLOYMENT

In order to validate and evaluate the design choices of the proposed solution, the system has been implemented and deployed at two critical infrastructures in the city of Santander.

A. "MERCADO DEL ESTE" DEPLOYMENT

A first deployment was done in one of the main markets in the city center of Santander. This deployment is intended to monitor the visitor's behavior in the "Mercado del Este" market, an old building restored in 2000 that contains shops, restaurants, a museum and a tourist office. This is a symmetric 60 × 40 meters building with three entrances in each of the longer facades. Figure 2 sketches the building where the system was deployed.

There are 8 measurement devices which uses the Raspberry Pi Model B+ board to support both, the End-Device and the Aggregator Device. The Raspberry Pi features a 900 MHz Quad-core ARM Cortex-A7 with 1 GB of RAM. Each device is equipped with a TP-LINK TL-WN722N USB adaptor with a 4 dBi onmidirectional antenna. This USB adaptor features the Atheros AR9271 chipset. In addition to the Wi-Fi adaptor, the devices also includes a Bluetooth 4.0 adaptor, and a DHT22 sensor, which can monitor temperature and humidity.

Taking advantage of having an indoor deployment, devices were powered through Power Over Ethernet (PoE), which transports data and electric power through conventional twisted pair Ethernet cabling.

B. AIRPORT DEPLOYMENT

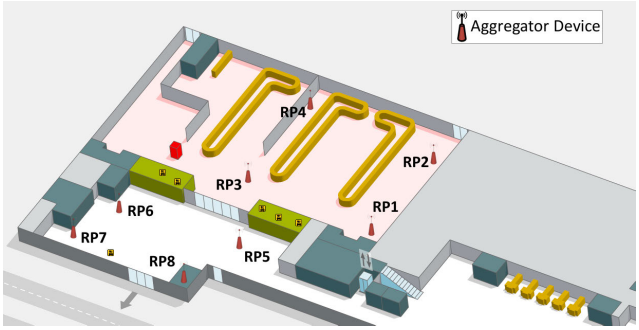

FIGURE 3. Airport deployment.

The second deployment was carried out in the airport to have an additional deployment in one of the most critical infrastructures in the city. This deployment was focused only on some of the public areas of the airport. Moreover, it has an additional constraint imposed by the airport security policy that does not allow connecting experimental devices to the airport's private wired network. This was addressed including a 3G USB adaptor to the routers used, two Mikrotik hEX PoE. Additionally, the Raspberry Pi used in this deployment were the Model 3 B+, featuring a Wi-Fi interface integrated in the board, and a new upgraded processor, 1.4 GHz Quad-core ARM Cortex-A53. Devices are also fed through PoE.

The deployment area is depicted in Figure 3. It covers two of the most crowded areas from the airport: the baggage reclaim area and the arrivals hall. It gives us the opportunity to monitor the crowd concentration of people in different periods, including stationary concentrations while waiting for luggage.

C. SOFTWARE IMPLEMENTATION

The implementation of the components running at the Aggregator-tier Devices have been done with Python. The Frame Feature Extraction and the Privacy Guarantor are both developed within the same multi-process Python script. Mainly, the Pyshark module has been employed for the capturing and filtering of the Wi-Fi frames, apart from the cryptographic libraries to execute the HMAC SHA-256 function. The implementation of the Sensor Broker and the Data Processor was carried out using Python scripts and the Scikit-learn [37] library. The database used for data persistence is built on top of a MySQL instance.

All the communications between the Aggregator-tier devices and the Processing-tier are secured. Using secured HTTPS web service to exchange data would not be appropriate in our setup, where data generation rate is very high, as HTTPS might produce a bottleneck. Raw data streaming channels might be better choice for this type of communication. However, the implementation can take a lot of effort, and it can be quite time-consuming to set them up properly. Alternatively, several transport communication frameworks are available to ease the usage of TCP streams. Among these alternatives, we can find RabbitMQ, Kafka, AMQP, which have a centralized broker; or ZeroMQ [38], which is brokerless. For our implementation, we chose ZeroMQ as it is the fastest alternative. More specifically, the ironhouse [39] implementation of ZMQ was employed, which provides out-of-the-box elliptic curve cryptography support.

Finally, some modifications were performed to the Raspberry Pi default configuration to ensure a robust behavior in a long-term scenario. Among them, the key adaptation was to make the file system read-only. This makes the Pi to work exclusively on RAM, and prevents SD card errors, thus no SD card replacement would be necessary in a long-term scenario. It has some drawbacks like the loss of persistence for some tasks (e.g. logs), but it enhances the security of the system as no information can be accessed if the deployed devices are tampered.

VI. A PROCESSING LAYER CROWD MANAGEMENT IMPLEMENTATION

The analysis of Wi-Fi frames to gather information of existing crowd concentration presents several issues that need to be addressed. In this sense, Wi-Fi signals behavior are not easy to predict as propagation depends on multiple factors, such as reflections, signal strength fluctuation, etc. Hence, detected signal strength is heavily dependent on the specifics. There

is not a single model that could fit any scenario. Thus, it is necessary to implement specific models depending on the area where the devices are deployed.

Deterministic solutions could provide a valid solution, for instance using thresholds on the RSSI to distinguish whether a device is within a specific area or not. However, this kind of solution is really bound to each deployment and require a thorough analysis of the environment, which makes these solutions hardly scalable.

For our implementation, we have chosen two probabilistic classifiers, the Logistic Regression and the Naive Bayesian classifiers, to assess the location of the detected devices. Additionally, we have used the Random Forest ensemble learning method limited to 10 trees to reduce the computational cost. The use of basic ML techniques instead of more complex AI mechanisms is motivated by three main aspects: (1) need for fast convergence. The Data Processor has to continuously update the crowd assessment based on the constant stream of frames captured on the field; (2) low computational complexity. Ideally, the Processing-tier would be deployed at edge servers to avoid unnecessary delays. Thus, only limited resources should be available; and (3) relaxed accuracy need. Exact estimation of the crowd is not mandatory for most of the crowd management applications that would consume the data provided by this system.

A. LOGISTIC REGRESSION CLASSIFIER

The Logistic Regression Classifier (LRC) models the probability that a certain input (X) belongs to an specific class (Y). It is an evolution of the traditional linear regression model where the output is a discrete categorical value.

In our case, considering that this classifier can estimate dichotomous and polytomous dependent variables, it was used to estimate the location of the detected devices within the deployment area (e.g. whether a detected device is within a building or not).

WE consider ($P(Y)$) as the probability of a certain event (Y), and that the Logit function can be expressed as a linear regression of independent variables as shown in Equation 1:

$$logit(P) = ln\left(\frac{P}{1-P}\right) = \theta_0 + \sum_{n=0}^{\infty} \theta_n X_n \tag{1}$$

where n is the number of dependent variables. From Equation 1 we can derive the probability of the different outputs:

$$P(Y) = \frac{1}{1 + e^{\theta_0 + \sum_{n=0}^{\infty} \theta_n X_n}} \tag{2}$$

If we apply the current deployment to Equation 2 we have the θ_n constants depending on the devices deployed to monitor the area. Similarly, the number of categories for the dependent variable (Y) will have to meet the number of areas we want to monitor (e.g. two if we want to assess whether the detected device is outside or inside the monitored area). Therefore, depending on the input, the model will assign a specific probability to each of the areas.

B. NAIVE BAYESIAN CLASSIFIER

Similarly, the Naive Bayesian Classifier (NBC) is a ML model that discriminates categorical outputs based on a set of features, assuming strong independence between them.

In our case, if we define X as the random variable representing the values obtained by the deployed devices, where n is the number of them, it is possible to estimate, applying the Bayes theorem, the probability of the smartphone to be in a specific area.

$$P(Y|X_1, X_2, \ldots X_n) \propto P(Y) \prod_{i=1}^{n} P(X_i|Y) \tag{3}$$

Equation 3 is a simplified version of the direct application of the Bayes theorem, in which proportionality is being considered, as the denominator does not change. Consequently, to find the area (Y) where the device is located, it is necessary to find the output Y with the biggest probability as defined in Equation 4.

$$Y = argmax_Y P(Y) \prod_{i=1}^{n} P(X_i|Y) \tag{4}$$

Hence, to create the model for the NBC, it is necessary to compute the probabilities obtained from a training dataset in the different areas.

C. RANDOM FOREST CLASSIFIER

The Random Forest Classiffier (RFC) is an ensemble learning method that is composed of multiple decision trees. This method was firstly presented by Ho [40] and extended by Breiman [41] lately. The main advantage of RFC over Decision Trees is the reduced variance while avoiding overfitting, thus improving the performance of the model. RFC uses recursive partitioning to generate the decision trees and aggregate the results. Decision trees are created by sampling the dataset with replacement to obtain a bootstrap sampling [42]. Therefore, the distribution of the samples are statistically identical in regards to the original dataset. Considering ($P(Y)$) as the probability of a certain event (Y) provided by each decision tree, we obtain the random forest probability by averaging all of them, as shown in 5, being D the number of Decision Trees used to model the RFC.

$$P'(Y) = \frac{1}{D} \sum_{n=0}^{D} P(Y) \tag{5}$$

Hence, the output Y will be the one with the biggest probability as shown in the equation 6.

$$Y = argmax_Y P'(Y) \tag{6}$$

D. TEST SCENARIO SETUP AND PERFORMANCE OF SELECTED CLASSIFIERS

In order to test the probabilistic classifiers we have performed several measurement campaigns, obtaining 246,500 and 1,540,168 Probe Request frames from both, the ''Mercado del

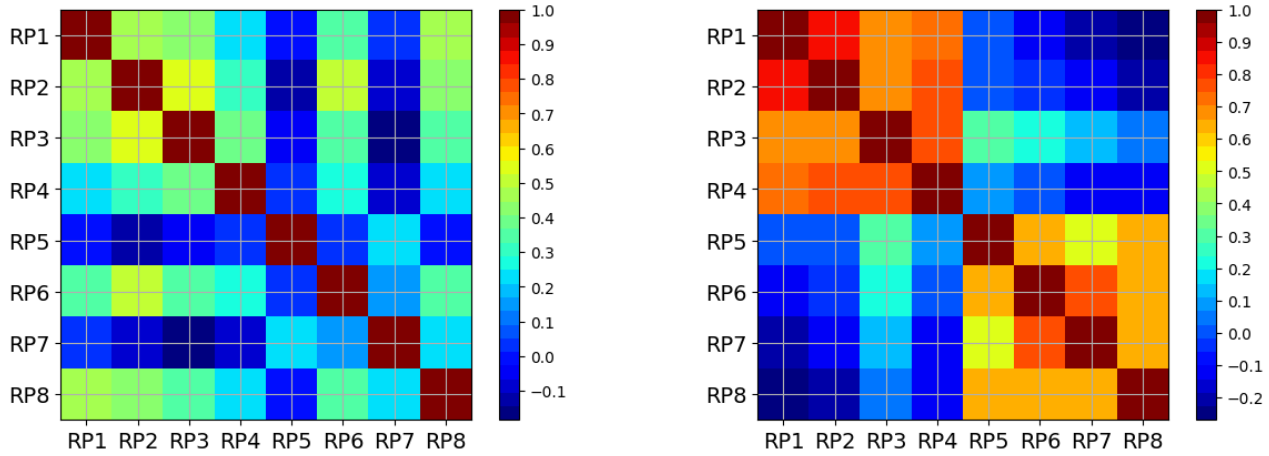

FIGURE 4. Pearson correlation matrix from "Mercado del Este" (left) and airport (right) deployments.

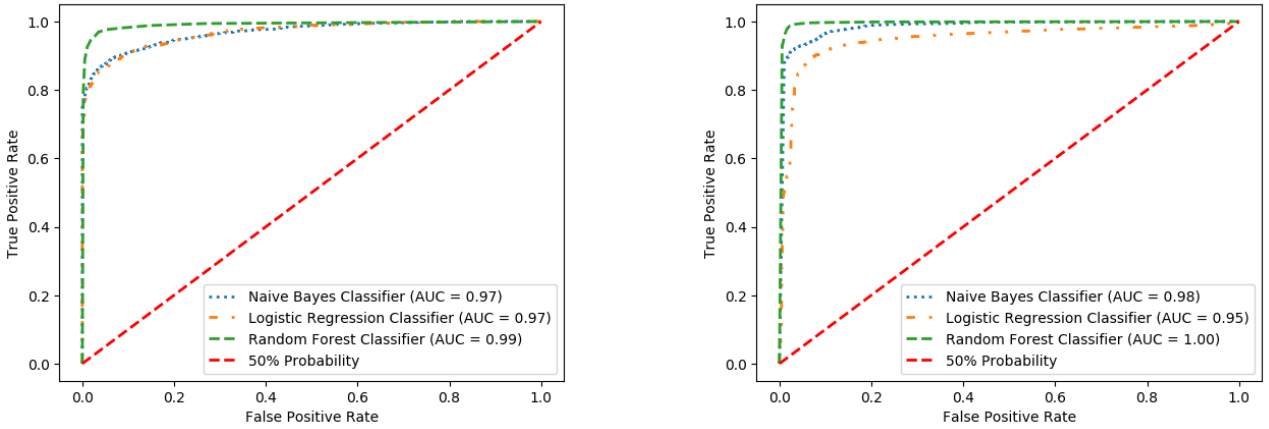

FIGURE 5. ROC curve for selected classifiers from "Mercado del Este" (left) and airport (right) deployments.

Este'' and Airport deployment, respectively. These measurement campaigns have been carried out using a set of 12 different smartphones, including Android, IOS and Windows Phone based operating systems, which were forced to send Probe Requests continuously. Furthermore, the measurements were taken in different spots evenly distributed, both inside and outside the buildings, with the ground truth annotation of the exact position and the time spent in each of the spots. Therefore, for the assessment of the classifiers accuracy, it is possible to match the frames sent from a specific location with their RSSI fingerprint.

For the analysis, the RSSI values gauged by the 8 measurement devices every one second from the same phone have been aggregated in a single vector. In order to avoid having blanks in these vectors, in case any of the measurement devices failed to get the RSSI, the corresponding element of the vector is set to -102 dBm. We are assuming that this kind of situations are due to not receiving the corresponding Probe Request. Hence, we are using 3dB less than the sensitivity of the Wi-Fi chipset used in the measurement devices. Finally, columns with less than two real RSSI values (i.e. not forced to the fixed value) were discarded.

Chosen probabilisic classifiers perform better with a low correlation among input variables. Therefore, we have studied the product-moment correlation coefficient from the measurements. As it is shown in Figure 4, there is strong relation among the values obtained by closer measurement devices. This is especially evident in the Airport deployment, where the correlation between measurement devices located in the same room (i.e. RP1-RP4 and RP5-RP8) are above 0.6. Both scenarios show a weak negative correlation between measurement devices that are far away from each other (i.e RP3 and RP7, RP1 and RP8). However, the overall situation is varied enough, remarkably at the ''Mercado del Este'' deployment, where correlation is generally below 0.4).

Half of the dataset was used for creating the models while the other half was used for the assessment. Figure 5 shows the Receiver Operating Characteristic (ROC) curves for both deployments. Attending to the Area Under the Curve (AUC) parameter, we can see that all the classifiers have a very good

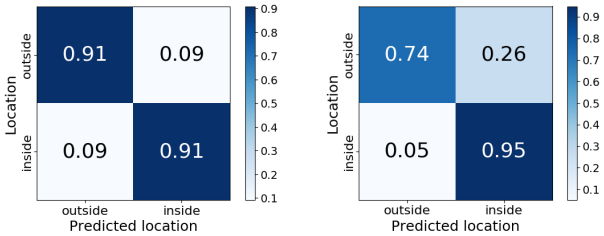

FIGURE 6. Confusion matrix using the logistic regression classifier in "Mercado del Este"" (left) and airport (right) deployments.

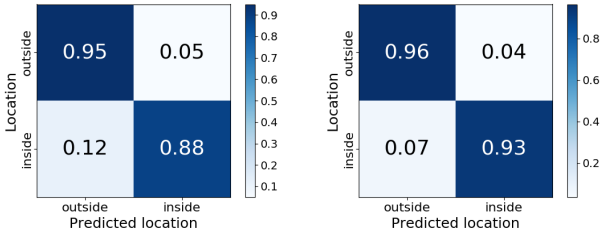

FIGURE 7. Confusion matrix using the Naive Bayes classifier in "Mercado del Este" (left) and airport (right) deployments.

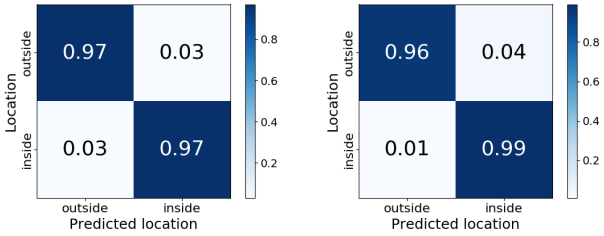

FIGURE 8. Confusion matrix using the random forest classifier in "Mercado del Este" (left) and airport (right) deployments.

performance for this classification problem. This result is endorsed by the confusion matrices shown in Figure 8. These matrices have resulted from applying both classifiers to the different deployments, where the decision for the predicted location is made when $P(Y) > 0.5$. As shown in the results, the RFC provides the best results, followed closely by NBC, which performs slightly better than the LRC. However, even the worst results (obtained by the LRC at the Airport) are quite satisfactory for this use case. This poorer performance of the LRC at the airport is mainly due to the location of the deployed devices that causes higher correlations. Finally, convergence times for both methods were short enough to guarantee real-time decisions.

All in all, the results of the analysis performed are good enough to validate our design premises as they demonstrate that using lightweight AI classifiers, which can easily fit with the capacities of any edge server, on top of available data from WiFi traces, it is possible to obtain really accurate estimations of crowds in public buildings.

VII. CONCLUSION

Crowd control in urban areas is becoming a major objective for smart cities. This paper have introduced a baseline AI-aided IoT-based architecture for crowd management systems taking advantage of Wi-Fi frames analysis. The presented architecture has a strong focus on privacy, addressing the restrictive European Union privacy regulation, which considers MAC addresses as personal datum for such types of deployments. Moreover, its design is ready to make the most of the fog-computing paradigm for enabling low latency decision making and for reducing its vulnerability to network bottlenecks.

Likewise, the paper describes the two deployments that have been carried out in the city of Santander, meant to validate and evaluate the solution in real-world situations. In this paper we have described the results of applying three different AI classifiers for crowd assessment to both deployments. The three mechanisms analysed have shown a good performance in deciding whether the targets were inside or outside the building. It is important to highlight that they were chosen because of their lightweight computational footprint and fast convergence time. This feature makes them good candidates to employ fog-computing approaches, thus reducing even more the decision time and satisfying, in a larger extend, the key design considerations described in the paper. In this sense, as it has been stated, the contributions of the paper are focused on the design, implementation, and subsequent validation, of the crowd-management system in a real-world scenario under the aforementioned design considerations of stringent privacy restrictions and lightness. Rather than researching on the most accurate crowd estimation approach, the paper has focused on the implementation of a crowd monitoring system addressing key requirements that are overseen in other studies and its validation through the integration and analysis of several lightweight classifiers, despite the fact that there are other analogous algorithms with similar or even better performance.

As part of the future work, we intend to include other classification techniques, including more computational-demanding algorithms, to evaluate which provides the best performance for positioning tasks. It is also foreseen the use of more complex machine learning techniques to extract additional features from the datasets, such as the prediction of periods were the monitored areas are most crowded. Furthermore, it is intended to study the implementation of unsupervised machine learning algorithms to assess crowds' behavior in critical infrastructures, with the inclusion of additional data sources (e.g. weather forecast).

ACKNOWLEDGEMENT

The authors would like to thank Mr. Sergio Ortega, from the Santander airport's authority, for his support.

REFERENCES

[1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[2] V. Cisco, "Cisco visual networking index: Forecast and trends, 2017–2022," Cisco, San Jose, CA, USA, White Paper, 2018.

[3] M. Mohammadi and A. Al-Fuqaha, "Enabling cognitive smart cities using big data and machine learning: Approaches and challenges," *IEEE Commun. Mag.*, vol. 56, no. 2, pp. 94–101, Feb. 2018.
[4] H. Hassanein, N. Zorba, S. Han, S. S. Kanhere, and M. Shukair, "Crowd management," *IEEE Commun. Mag.*, vol. 57, no. 4, pp. 18–19, Apr. 2019.
[5] U DESA, New York, NY, USA. (2018). *World Urbanization Prospects: The 2018 Revision, Key Facts*. Accessed: Dec. 20, 2018. [Online]. Available: https://population.un.org/wup/Publications/
[6] V. A. Sindagi and V. M. Patel, "A survey of recent advances in CNN-based single image crowd counting and density estimation," *Pattern Recognit. Lett.*, vol. 107, pp. 3–16, May 2018.
[7] A. B. Chan, Z.-S. John Liang, and N. Vasconcelos, "Privacy preserving crowd monitoring: Counting people without people models or tracking," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–7.
[8] F. Adib and D. Katabi, *See Through Walls With WiFi!*, vol. 43, no. 4. New York, NY, USA: ACM, 2013.
[9] T. Franke, P. Lukowicz, and U. Blanke, "Smart crowds in smart cities: Real life, city scale deployments of a smartphone based participatory crowd management platform," *J. Internet Services Appl.*, vol. 6, no. 1, p. 27, Aug. 2015.
[10] W. Kang and Y. Han, "SmartPDR: Smartphone-based pedestrian dead reckoning for indoor localization," *IEEE Sensors J.*, vol. 15, no. 5, pp. 2906–2916, May 2015.
[11] S. Qiu, Z. Wang, H. Zhao, K. Qin, Z. Li, and H. Hu, "Inertial/magnetic sensors based pedestrian dead reckoning by means of multi-sensor fusion," *Inf. Fusion*, vol. 39, pp. 108–119, Jan. 2018.
[12] J. Wang, Y. Wang, D. Zhang, L. Wang, C. Chen, J. W. Lee, and Y. He, "Real-time and generic queue time estimation based on mobile crowdsensing," *Frontiers Comput. Sci.*, vol. 11, no. 1, pp. 49–60, Feb. 2017.
[13] Y. Wang, J. Wang, and X. Zhang, "QTime: A queuing-time notification system based on participatory sensing data," in *Proc. IEEE 37th Annu. Comput. Softw. Appl. Conf.*, Jul. 2013, pp. 770–777.
[14] J. Steenbruggen, E. Tranos, and P. Nijkamp, "Data from mobile phone operators: A tool for smarter cities?" *Telecommun. Policy*, vol. 39, nos. 3–4, pp. 335–346, May 2015.
[15] F. Calabrese, L. Ferrari, and V. D. Blondel, "Urban sensing using mobile phone network data: A survey of research," *ACM Comput. Surv.*, vol. 47, no. 2, p. 25, Jan. 2015.
[16] A. Shahmansoori, G. E. Garcia, G. Destino, G. Seco-Granados, and H. Wymeersch, "Position and orientation estimation through millimeter-wave MIMO in 5G systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1822–1835, Mar. 2018.
[17] F. Zafari, A. Gkelias, and K. K. Leung, "A survey of indoor localization systems and technologies," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2568–2599, 3rd Quart., 2019.
[18] F. Aloul, A. Sagahyroon, A. Al-Shami, I. Al-Midfa, and R. Moutassem, "Using mobiles for on campus location tracking," in *Proc. 7th Int. Conf. Adv. Mobile Comput. Multimedia (MoMM)*. New York, NY, USA: ACM, 2009, pp. 231–235.
[19] F. Meneses and A. Moreira, "Large scale movement analysis from WiFi based location data," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Nov. 2012, pp. 1–9.
[20] S. Sendra, M. Garcia, C. Turro, and J. Lloret, "People mobility behaviour study in a university campus using WLANs," in *Proc. 3rd Int. Conf. Mobile Ubiquitous Comput., Syst., Services Technol.*, Oct. 2009, pp. 124–129.
[21] L. Vu, K. Nahrstedt, S. Retika, and I. Gupta, "Joint bluetooth/WiFi scanning framework for characterizing and leveraging people movement in university campus," in *Proc. 13th ACM Int. Conf. Modeling, Anal., Simulation Wireless Mobile Syst. (MSWIM)*. New York, NY, USA: ACM, 2010, pp. 257–265.
[22] S. Woo, S. Jeong, E. Mok, L. Xia, C. Choi, M. Pyeon, and J. Heo, "Application of WiFi-based indoor positioning system for labor tracking at construction sites: A case study in Guangzhou MTR," *Autom. Construct.*, vol. 20, no. 1, pp. 3–13, Jan. 2011.
[23] B. Zhou, Q. Li, Q. Mao, and W. Tu, "A robust crowdsourcing-based indoor localization system," *Sensors*, vol. 17, no. 4, p. 864, Apr. 2017.
[24] B. Bonne, A. Barzan, P. Quax, and W. Lamotte, "WiFiPi: Involuntary tracking of visitors at mass events," in *Proc. IEEE 14th Int. Symp. 'World Wireless, Mobile Multimedia Netw.' (WoWMoM)*, Jun. 2013, pp. 1–6.
[25] B. S. Ciftler, S. Dikmese, I. Guvenc, K. Akkaya, and A. Kadri, "Occupancy counting with burst and intermittent signals in smart buildings," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 724–735, Apr. 2018.
[26] S. Longo and B. Cheng, "Privacy preserving crowd estimation for safer cities," in *Proc. ACM Int. Joint Conf. Pervas. Ubiquitous Comput. Proc. ACM Int. Symp. Wearable Comput. (UbiComp)*. New York, NY, USA: ACM, 2015, pp. 1543–1550.
[27] L. Sanchez, L. Muñoz, J. A. Galache, P. Sotres, J. R. Santana, V. Gutierrez, R. Ramdhany, A. Gluhak, S. Krco, E. Theodoridis, and D. Pfisterer, "SmartSantander: IoT experimentation over a smart city testbed," *Comput. Netw.*, vol. 61, pp. 217–238, Mar. 2014.
[28] R. Van Kranenburg, N. Stembert, M. V. Moreno, A. F. Skarmeta, C. López, I. Elicegui, and L. Sánchez, "Co-creation as the key to a public, thriving, inclusive and meaningful EU IoT," in *Proc. Int. Conf. Ubiquitous Comput. Ambient Intell.* Cham, Switzerland: Springer, 2014, pp. 396–403.
[29] M. Tesanovic, E. Conil, A. De Domenico, R. Aguero, F. Freudenstein, L. M. Correia, S. Bories, L. Martens, P. M. Wiedemann, and J. Wiart, "The LEXNET project: Wireless networks and EMF: Paving the way for low-EMF networks of the future," *IEEE Veh. Technol. Mag.*, vol. 9, no. 2, pp. 20–28, Jun. 2014.
[30] R. Agarwal, D. G. Fernandez, T. Elsaleh, A. Gyrard, J. Lanza, L. Sanchez, N. Georgantas, and V. Issarny, "Unified IoT ontology to enable interoperability and federation of testbeds," in *Proc. IEEE 3rd World Forum Internet Things (WF-IoT)*, Dec. 2016, pp. 70–75.
[31] J. Lanza, L. Sanchez, J. R. Santana, R. Agarwal, N. Kefalakis, P. Grace, T. Elsaleh, M. Zhao, E. Tragos, H. Nguyen, F. Cirillo, R. Steinke, and J. Soldatos, "Experimentation as a service over semantically interoperable Internet of Things testbeds," *IEEE Access*, vol. 6, pp. 51607–51625, 2018.
[32] T. Akiyama, S. Murata, K. Tsuchiya, T. Yokoyama, M. Maggio, G. Ciulla, J. R. Santana, M. Zhao, J. B. D. Nascimento, and L. Gürgen, "FESTIVAL: Design and implementation of federated interoperable smart ICT services development and testing platform," *J. Inf. Process.*, vol. 25, pp. 278–287, Mar. 2017.
[33] G. Solmaz, F.-J. Wu, F. Cirillo, E. Kovacs, J. R. Santana, L. Sanchez, P. Sotres, and L. Munoz, "Toward understanding crowd mobility in smart cities through the Internet of Things," *IEEE Commun. Mag.*, vol. 57, no. 4, pp. 40–46, Apr. 2019.
[34] European Commission, "Regulation (EU) 2016/679 of the European parliament and of the council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46," *Off. J. Eur. Union*, vol. 59, nos. 1–88, p. 294, Apr. 2016.
[35] A. Pfitzmann and M. Köhntopp, "Anonymity, unobservability, and pseudonymity—A proposal for terminology," in *Designing Privacy Enhancing Technologies*. Berlin, Germany: Springer, 2001, pp. 1–9.
[36] Agencia Española de Protección de Datos. (Oct. 2016). *Orientaciones y Garantías en Los Procedimientos de Anonimización de Datos Personales*. Accessed: May 3, 2019. [Online]. Available: https://www.aepd.es/media/guias/guia-orientaciones-procedimientos-anonimizacion.pdf
[37] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Oct. 2011.
[38] Hintjens, Pieter and Sustrik, Martin. *ZeroMQ*. Accessed: May 3, 2019. [Online]. Available: http://zeromq.org/
[39] Hintjens, Pieter. *Ironhouse ZeroMQ*. Accessed: May 3, 2019. [Online]. Available: http://hintjens.com/blog:49
[40] T. Kam Ho, "Random decision forests," in *Proc. 3rd Int. Conf. Document Anal. Recognit.*, vol. 1, 1995, pp. 278–282.
[41] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[42] B. Efron, "Bootstrap methods: Another look at the jackknife," in *Breakthroughs in Statistics*. New York, NY, USA: Springer, 1992, pp. 569–593.

JUAN RAMÓN SANTANA is currently a Senior Research Fellow with the University of Cantabria, Spain. He has been involved in several Smart City international projects, from which we can highlight SmartSantander, where he carried out the integration and deployment of the SmartSantander communication infrastructure. Beyond SmartSantander, he has been involved in other EU projects, such as CLouT or FESTIVAL, as the Work Package Leader. He has authored more than 30 publications, including conferences, journals, and book chapters. His research interest includes wireless sensor networks (WSN) for the smart city.

LUIS SÁNCHEZ is currently an Associate Professor with the University of Cantabria. He has authored more than 80 papers at international journals and conferences and coauthored several books. He often participates in panels and round tables discussing about innovation supported by the IoT in smart cities. He acts as an Expert for reviewing and evaluating Research and Development proposals. His research interests include the IoT-enabled smart cities, meshed networking on heterogeneous wireless scenarios, and optimization of network performance through cognitive networking techniques.

PABLO SOTRES received the Telecommunications Engineering degree from the University of Cantabria, Spain, in 2008. He has been involved in several different international projects framed under the Smart City paradigm, such as SmartSantander; and related to inter-testbed federation, such as Fed4FIRE, Fed4FIRE+, the Wise-IoT, and FED4SAE. He is currently a Research Fellow with the Network Planning and Mobile Communications Laboratory, Communications Engineering Department, University of Cantabria.

JORGE LANZA received the Ph.D. degree in telecommunications engineering from the University of Cantabria, Spain, in 2014. He has participated in several research projects, national and international, with private and public funding. He is currently a Senior Researcher with the Network Planning and Mobile Communications Laboratory, University of Cantabria. His current research interests include the IoT infrastructures toward federating deployments in different locations using semantics technologies. In addition, his work has included combined mobility and security for the wireless Internet.

TOMÁS LLORENTE received the M.Sc. degree in telecommunication engineering from the University of Cantabria, in 2018. He is currently working as a Cybersecurity Consultant with the University of Cantabria. He has been involved in several research areas, including the analysis of a virtual lab applied to the cloud's vulnerabilities or the usage of the IoT infrastructure to determine people location in buildings. His research interests include study of cybersecurity threats in telecommunication networks and the Internet of Things.

LUIS MUÑOZ (Senior Member, IEEE) received the Telecommunications Engineering and Ph.D. degrees from the Polytechnical University of Cataluña, Spain, in 1990 and 1995, respectively. He is currently the Head of the Network Planning and Mobile Communications Laboratory, University of Cantabria, Spain. He has participated in several National and European research projects, and was the Technical Manager of SmartSantander. He has authored over 150 journal and conference papers. He also serves as a Consultant for the Spanish Government as well as for different companies in Europe and USA. His research interests include advanced data transmission techniques, heterogeneous wireless multihop networks, and applied mathematical methods for telecommunications.