sensors

Article

Crowd monitoring in smart tourism destinations based on WiFi scanning

Alberto Berenguer [1], David Fernández [2], Andrea Gómez-Oliva [2], Josep A. Ivars-Baidal [3], Antonio J. Jara [4,*], Jaime Laborda [5], Jose-Norberto Mazón [1] and Angel Perles [5]

Citation: Lastname, F.; Lastname, F.; Lastname, F. Title. *Sensors* **2021**, *21*, x. https://doi.org/10.3390/xxxxx

Academic Editor: Firstname Lastname
Received: date
Accepted: date
Published: date

[1] Instituto Universitario de Investigación Informática, Universidad de Alicante, Spain; aberenguer@dlsi.ua.es, jnmazon@ua.es
[2] HOP Ubiquitous S.L., Calle Luis Buñuel No. 6, 30562 Ceutí, Murcia, Spain; andrea@hopu.org, davidfr@hopu.org
[3] Instituto Universitario de Investigaciones Turísticas, Universidad de Alicante, Spain; josep.ivars@ua.es
[4] Institute of Information Systems, University of Applied Sciences Western Switzerland, ConEx Lab, 3960 Sierre, Switzerland; jara@ieee.org
[5] ITACA Institute, Universitat Politècnica de València, Camino de Vera, s/n 46022 Valencia, Spain; jlaborda@itaca.upv.es, aperles@disca.upv.es
* Correspondence: jara@ieee.org

Abstract: Crowd monitoring has been an essential measure to deal with overtourism problems in urban destinations in the pre-Covid era, and it is going to play a key role in the pandemic scenario to restart tourism and make destinations safer. Importantly, a Destination Management Organization (DMO) of a smart tourism destination needs to deploy a technological layer for crowd monitoring that allows data gathering for counting visitors and distinguishing them from residents. The correct identification of visitors versus residents by a DMO, while privacy rights (e.g., Regulation EU 2016/679, also known as GDPR) are ensured, is an open problem that has not been fully solved. In this paper, we describe a novel approach to get crowd data by processing (i) massive scanning of WiFi access points of the smart tourism destination to find SSIDs, as well as (ii) the exposed Preferred Network List (PNL) contained the SSIDs of WiFi access points that WiFi-enabled mobile devices are likely to connect. This data allows us to provide the number of visitors and residents of a crowd in a given point of interest of a smart tourism destination. A pilot study has been conducted in the city of Alcoi (Spain) comparing data coming from our approach with data provided by manually filled surveys from the Alcoi Tourist Info office, thus showing the feasibility of our approach to enrich the information system of a smart tourism destination.

Keywords: smart tourism destination; GDPR; crowd monitoring; WiFi scanning

1. Introduction

The evolution of information and communications technology (ICT) has been a major factor of disruption in the tourism industry. Recently, the smart tourism concept has been coined as a distinct step in the evolutionary relationship of ICT and tourism, characterized by the integration of the physical and the digital world [1]. Smart ecosystems have emerged as tourism systems that take advantage of novel technologies (such as IoT) and the intensive use of information in creating, managing and delivering intelligent touristic services/experiences [2]. From the management point of view, a new scenario appears under the smart destination approach [3], that contributes to improve sustainability, competitiveness and tourism experiences [4].

The generation of new information systems is one of the main pillars of smart tourism destinations. These systems include market intelligence linked to the different phases of the travel life cycle (inspiration, booking, experiencing, and sharing) and, specifically, the monitoring of tourist behavior at destination, as key information for advanced management. The Destination Management Organization (DMO) should function as a smart hub that coordinates all relevant sources of information and makes it accessible for different users [5]. In the case of the tourist behaviour monitoring, the DMO must deploy a technology layer that allows data gathering for different purposes.

In this context, measuring the spatial behavior of tourism is a challenge for both researchers and practitioners, above all tracking visitors movements and their concentration at particular points of interest [6]. Actually, crowd monitoring has been an essential measure to deal with overtourism problems in urban destinations in the pre-Covid era [7] and it is going to play a key role in the pandemic scenario to restart tourism and make destinations safer [8]. Therefore, the DMO must have enough and proper crowd monitoring data to make informed decisions and offer visitors a secure experience, while residents keep safe as well, e.g. avoiding overcrowding of visitors.

However, crowd monitoring by the DMO has not been widely adopted due to the need to associate people with some tracking technologies, while fulfilling the privacy restrictions imposed by regional legislation. It should be noted that this work is geographically framed within the European Community, so the Regulation (EU) 2016/679 [9] applies. This regulation is better known as the General Data Protection Regulation (GDPR) and seeks to harmonize data privacy laws across Europe. This is one of the most restrictive legislation in the world in terms of safeguarding the privacy of citizens. Therefore, it is a great challenge for an effective European DMO to comply with such regulation. Quite the contrary, most efforts in crowd monitoring have been done by using image processing, hindering privacy regulations enforcement [10]. On the other hand, mobile phone operators and companies such as Google and Apple know all our movements in real-time (and, in many cases, our interests) but, paradoxically, the exploitation of this information is forbidden to the DMO in order to comply with the GDPR. On the other hand, it is also important for the DMO to avoid total dependence on data providers such as telecommunication companies or booking platforms. Despite restrictions on widespread access to information collected by mobile phones, DMOs have devised ways to circumvent legal restrictions by promoting user consent, directly or indirectly, to share information. The most commonly used approaches are the use of destination apps, QR codes, and NFC tags [3,6]. Unfortunately, it is difficult to convince tourists to install destination apps and they are unlikely to make use of the QR or NFC tags made available by tourism services.

To overcome these shortcomings, WiFi-based approaches have emerged lately as a promising way to perform crowd monitoring, since they better fit with privacy regulations [10]. As a matter of fact, WiFi scanning of mobile phones with WiFi interface enabled is able to capture the unique MAC (Media Access Control) address of the devices to track and only to count the number of people in a certain space to avoid incurring privacy issues. A prominent open project in this regard is [11]. In this sense, techniques for estimating the number of mobile devices present at a certain location and time, through analysis of WiFi probe requests from smart devices have been proposed so far (e.g., [12]). Interestingly, other information also exposed by WiFi-enabled mobile devices via probe requests is the list of its preferred WiFi access points (the Preferred Network List, aka PNL) in the form of SSIDs (Service Set Identifier).

The research question that guides this work is: Would it be feasible to exploit this WiFi scanning information to go beyond detecting the number of persons at a certain location and time, and distinguishing visitors from residents in a smart tourism destination? And, of course, being compliant with the GDPR regulation. As stated by [13], this is still considered an open problem for smart tourism destinations, and it is also an initial step to further consider tourist digital footprints or data traces from tourist activities (since they occur if a person can be considered a tourist) [13].

To this aim, the main contribution of this paper is an approach to (i) get crowd data by processing smartphone device signals based on WiFi scanning, as well as (ii) differentiate the percentage of visitors and residents in a smart tourism destination. Our approach is based on the mass collection of local SSIDs and comparison with the preferred WiFi network information exposed by mobile terminals. In order to comply with GDPR, encryption and data locality mechanisms have been devised to ensure privacy. A pilot study has been conducted in the city of Alcoi (Spain) comparing data coming from our approach with data provided by manually filled surveys from the Alcoi Tourist Info office. First results show that our approach is a viable solution that could help the DMO to make informed decisions and offer visitors a secure experience, while residents keep safe as well, avoiding overcrowding of people in the pandemic scenario.

The article is structured as follows: Section 2 presents our crowd monitoring approach for detecting visitors and residents of a smart tourism destination; Section 3 shows the results of a pilot study in the city of Alcoi (Spain). Finally, the discussion about the obtained results and future work are presented in Section 4.

2. Visitor crowd monitoring approach

2.1 Architecture

The aim of our approach is to collect and process data from mobile devices to estimate the amount of people visiting a destination (i.e., visitors) in relation to the number of residents, for each day of the year. To do this, our visitor crowd monitoring approach takes advantage of the fact that smartphone devices with enabled WiFi interface periodically scan nearby WiFi access points available for connection [12]. This is done by sending (regardless the device is connected to a WiFi access point) control frames, named "probe requests". Information in this probe request contains, among others, the Preferred Network List, which contains the identifier of the WiFi access points (Service Set Identifier or SSID) to which the device has already connected to [12]. Therefore, our visitor crowd monitoring approach identifies those mobile devices whose WiFi SSID stored as preferred for connection is located in the destination (i.e., residents) and differentiate them from those devices whose preferred SSID is not from the destination (i.e., visitors).

Figure 1 shows a diagram of the proposed architecture, whose main elements are described next:

- A collector of SSIDs of the existing WiFi access points in a specific location (e.g., a city). This collector aims to create an accurate database of the available SSIDs in a tourist destination (top left of the figure).
- A collector of SSIDs of preferred WiFi access points coming from probe request data frames of mobile devices detected by HOPU Smart Spots [14] (top right of the figure).
- A postprocessing cloud infrastructure to determine the daily amount of visitors in different locations around the destination (bottom of the figure).

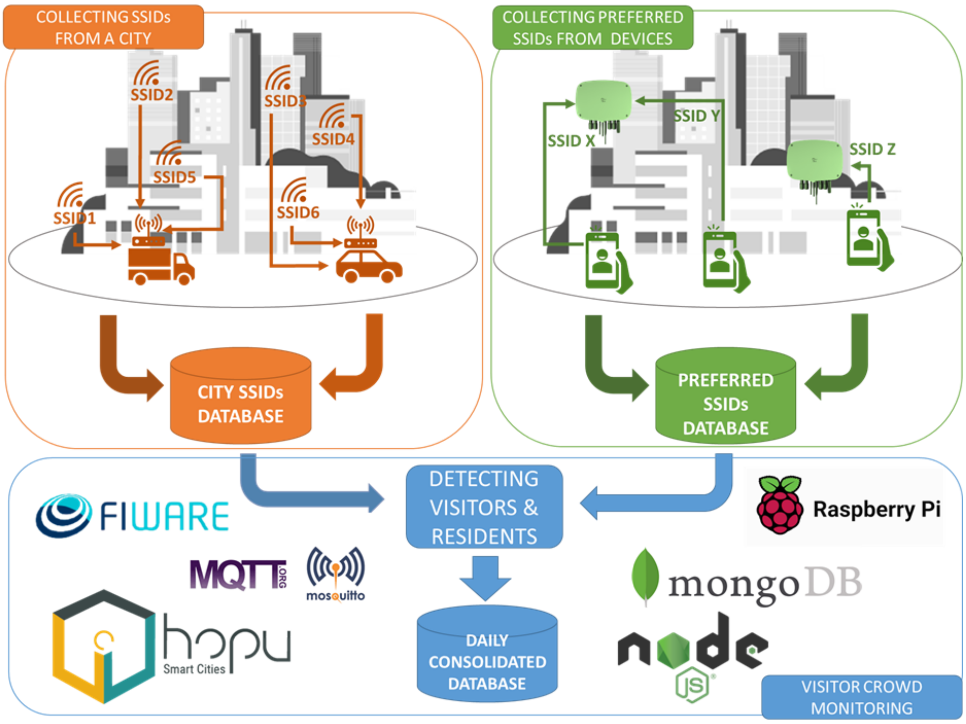


Figure 1. Architecture of our visitor crowd monitoring approach.

In the following subsections, each of the parts of the proposed approach is described in detail.

2.2. Collection of SSIDs from a smart tourism destination

WiFi access points usually expose their SSIDs by using periodic beacon frames. This is a periodic advertisement to inform any listening devices that this SSID is available and has particular features. Client devices depend upon these beacon frames to discover which networks are available (passive scanning), and to ensure that the networks that they are associated with are actually still present and available.

These SSIDs can be 0 to 32 bytes long and are, in general, a natural language string of arbitrary characters. An analysis of the exposed patterns worldwide shows that these strings tend to be locally identifiable and unique.

Collecting SSIDs is fairly immediate and has been massively employed by companies such as Google, Mozilla, etc. to improve their mobile device geolocation systems. A fairly comprehensive table of companies that collect this information for the so named WPS (WiFi Positioning System) and its availability can be found in [15]. Capturing SSIDs is not exempt from controversy, being remarkable the scandal around the surreptitious gathering of WiFi data while capturing video footage and mapping data for Google's Street View service [16]. As a result of this scandal, proposals such adding the tag "opt_out" to the SSIDs has been utilized for opting out to the capture of this service.

The legality of whether or not this information can be collected, especially in the European space, is unclear. In our particular case, the project is part of a public municipal service (city of Alcoi in Spain), so extreme measures have been taken in terms of legality and privacy, in agreement with the legal advisory services of the municipality. Basically, the approach followed is to immediately encrypt each of the captured SSID and store only the timestamps and the city where it was captured (no more specific location is required).

Although the most immediate implementation of the SSID capture system would be the use of a mobile phone application such as Wigle one [17], it was decided to develop a custom SSID collector to simplify the development and to improve the performance of SSID frame capture using high gain antennas.

To test the proposal, the set-up shown in Figure 2 was used. The key elements of this set-up are:

- A Raspberry Pi 4 Model B computer. It is a popular low-cost computer capable of running a full Linux operating system and thus allowing the development of complex applications.
- A dual-band USB WiFi dongle with high-gain antenna model TP-Link Archer T3U Plus with a sensitivity of about -75 dBm for 2.4 Ghz band and -70 dBm for 5 GHz band.
- An Anker PowerCore Essential 20000 PD powerbank to provide energy to the Raspberry Pi computer during journeys around the city. It should be noted that the Raspberry Pi 4 model used has a non-standard power supply design that makes it difficult to use a conventional powerbank.

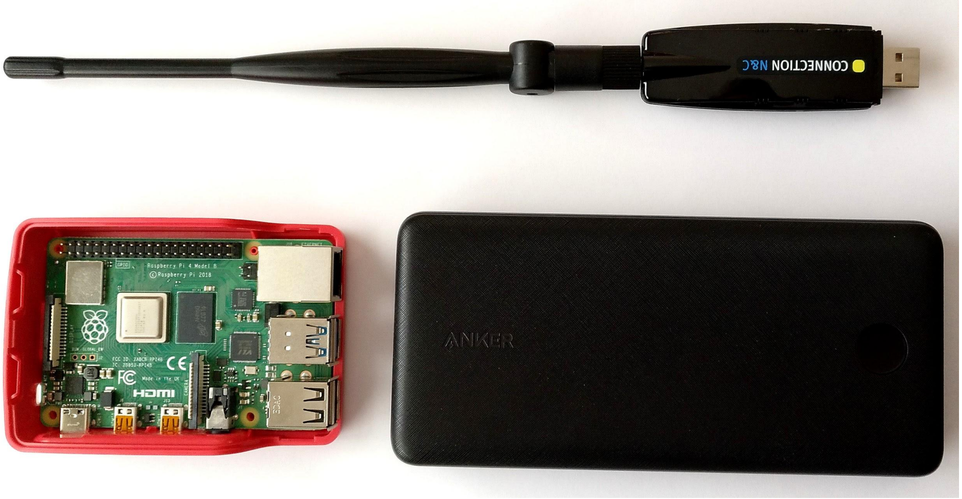


Figure 2. Key hardware components of the SSID city capture system: Raspberry Pi 4 model B minicomputer, Anker powerbank, USB WiFi dongle with high-gain antenna.

The antenna type election for the SSIDs scanner has been a key element because we intend to capture SSIDs of adjacent buildings from the ground level. In this sense, we have selected an omnidirectional antenna with a gain of about 5dBi to offer a reasonable compromise between range and an adequate reception pattern to pick up the SSIDs of the WiFi routers installed in the adjacent buildings.

Raspberry Pi collects SSID data from WiFi networks in a city and SHA1 hash algorithm is applied to the SSID of any encountered Wi-Fi access point, as well as the date of capture in timestamp format and the name of the city where it occurred.

Specifically, we have developed two software components:

- The first one is focused on collecting all the near availables SSIDs. It aims to select the network interface used to capture the data and set a refresh time (the time between capture of each SSID). Once configured this software executes a script that returns the available SSIDs and then cleans the output and applies a SHA1 hash encoding to ensure the privacy of the data. Finally, the collected data is written into a flat file in a SD memory. This software is running as a service, which means that if the Raspberry is on and the wireless card is connected, it will be collecting data.
- The second part takes care of the exportation of the data from the Raspberry to a cloud database described later. Database availability is checked before stopping the service running the SSID collection. It sets the date and the location, and then reads the file generated by the other script, thus inserting each SSID saved with a current date and the location. Finally, it checks if the quantity of the uploaded files corresponds to the number of SSIDs which appear on the text file to ensure that all the data collected is exported to the server. If everything is correct, the text file is deleted and the service of the collection script restarted.

2.3. Collection of preferred SSIDs from mobile devices

The collection of preferred SSIDs is done by a modified version of the HOPU Smart Spot device shown in Figure 3. This is a multi-sensor IoT device devoted to smart-city applications. In our case, we benefit from its WiFi crowd monitoring capability that has been adequately adapted to achieve the desired characteristics described below.

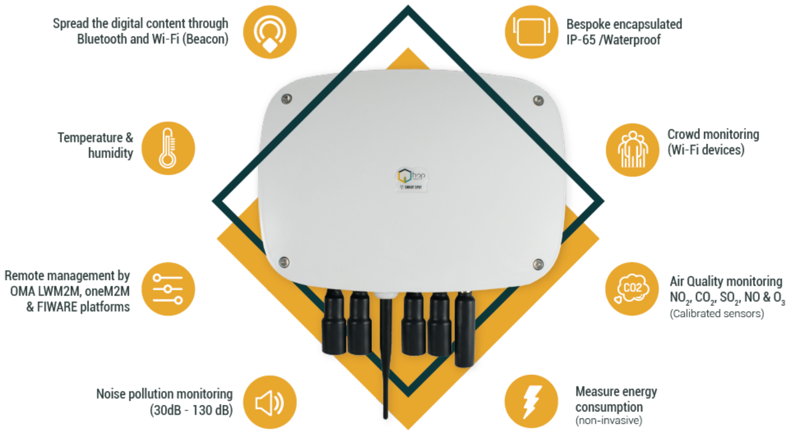

Spread the digital content through Bluetooth and Wi-Fi (Beacon)
Bespoke encapsulated IP-65 /Waterproof
Temperature & humidity
Crowd monitoring (Wi-Fi devices)
Remote management by OMA LWM2M, oneM2M & FIWARE platforms
Air Quality monitoring NO_2, CO_2, SO_2, NO & O_3 (Calibrated sensors)
CO2
Noise pollution monitoring (30dB - 130 dB)
Measure energy consumption (non-invasive)

Figure 3. HOPU Smart Spot IoT device.

The HOPU Smart Spot device listens for “probe request” packets sent by mobile devices. These packets correspond to those sent by mobile phones when searching for a WiFi access point to connect to. Normally the search for a WiFi access point may be accompanied by the name of the particular WiFi access point being searched for. Once a mobile device has been detected, within HOPU Smart Spot itself, a SHA1 hash algorithm is applied to the MAC and SSID (from the Preferred Network List or PNL) detected, making it impossible to obtain the MAC and SSID back again, thus anonymizing the user's information on the device itself after it has been received.

MACs and SSIDs anonymized using the SHA1 hash are stored in volatile memory and counted in 1, 5, and 10-minute intervals internally within the device. MACs (and corresponding SSID from the PNL) are discarded/deleted from memory when they are not detected by the device and are outside the maximum time range (10 minutes).

Every minute, the HOPU Smart Spot device counts the devices detected in each of the intervals (1, 5, and 10 minutes) and sends this information through the integration protocols available in the device (LwM2M/MQTT/LoRa/SENTILO/FIWARE).

This whole process of hashing MACs and SSIDs within the device itself and not storing them in non-volatile memory is carried out in order to protect citizens' data in accordance with the current data GDPR protection law.

It should be noted that the SSID collection is performed by means of hardware devices where no data is stored. On the other hand, the technologies used are innovative but mature enough to avoid any security risk. Specifically, FIWARE[1] framework is considered in order to use the MQTT protocol for message broker through the Mosquitto implementation, and MongoDB databases and servers that meet all security requirements according to European standards.

2.3. Cloud-based data storage and analysis infrastructure

To store and analyse the collected data, an Ubuntu server virtual machine has been deployed in a cloud infrastructure. In this machine, a MongoDB database has been created containing four collections, namely *ssidCollect*, *crowdLevel*, *places*, and *dataSSIDCollector*. The purpose of these collections are:

- *ssidcollect*: contains the data coming from each sensor of the HOPU Smart Spot devices. Specifically, this collection contains the following attributes:
    - *sensor* (string): it stores the id of the sensor from which it has obtained the data.
    - *date* (timestamp): the exact date on which the information has been captured.
    - *mac* (SHA1): information obtained by the sensor, specifically the MAC of the device after applying the SHA1 algorithm.
    - *ssid* (SHA1): information obtained by the sensor, specifically the SSID of the PNL from the corresponding MAC, or rather the summary function of that SSID after applying the SHA1 algorithm.
- *crowdLevel*: stores the number of people in a place at a given time. The following attributes are stored in this collection:
    - *sensor* (string): stores the id of the sensor that has obtained the data.
    - *date* (timestamp): the exact date on which the information has been captured.
    - *crowd*: stores the number of people detected by the sensor nearby.
- *places*: stores information about the place where the sensors are installed. Specifically, this collection contains the following attributes:
    - *sensor* (string): stores the id of the sensor that has obtained the information.
    - *name* (string): name of the place where the sensor is located.
    - *description* (string): a short description of the sensor location.
    - *coords* (coordinates): coordinates of the sensor location.
    - *city* (string): name of the city where the sensor is located.
- *dataSSIDCollector*: contains the SSID data collected by the Raspberry Pi. Attributes are as follows:
    - *date* (timestamp): exact date on which the data has been captured.

[1] https://www.fiware.org/

○ *ssid* (SHA1): information obtained by the Raspberry , specifically the SSID, or rather the summary function of that SSID after applying the SHA1 algorithm.
○ *location* (string): name of the city where the data was obtained.

Also, the collection *summaryInfo* is created to store daily aggregated data. This collection has the summary of the data captured by the sensor for each day (including the number of visitors and residents). This collection contains the following attributes:

○ *sensor* (string): stores the id of the sensor that has obtained the information.
○ *date* (timestamp): day for which the summary information is displayed.
○ *nVisitors* (integer): number of visitors detected.
○ *nResidents* (integer): number of residents detected.
○ *total* (integer): number of total people detected.

Collection *summaryInfo* is filled by applying a consolidation algorithm, whose pseudocode is next shown. This algorithm aims at using the collected data to obtain a daily summary of the number of visitors and residents detected by a specific sensor.

The *SSIDRatio* is the percentage of preferred SSIDs of a device that must belong to the city to be considered as a device owned by a resident. Default value is 0.5. On the other side, *frequentMAC* is the number of times that a MAC is detected each day for the same sensor in order to be counted as a unique device (e.g., because it belongs to a worker). Default value is 50. Finally, *list_frequent_SSIDs* is a list whose elements are hashed SSIDs. They come from a data source with common SSID names that cannot be classified. Data come from https://wigle.net/csv/ssid.csv and the reason for this usage is explained later.

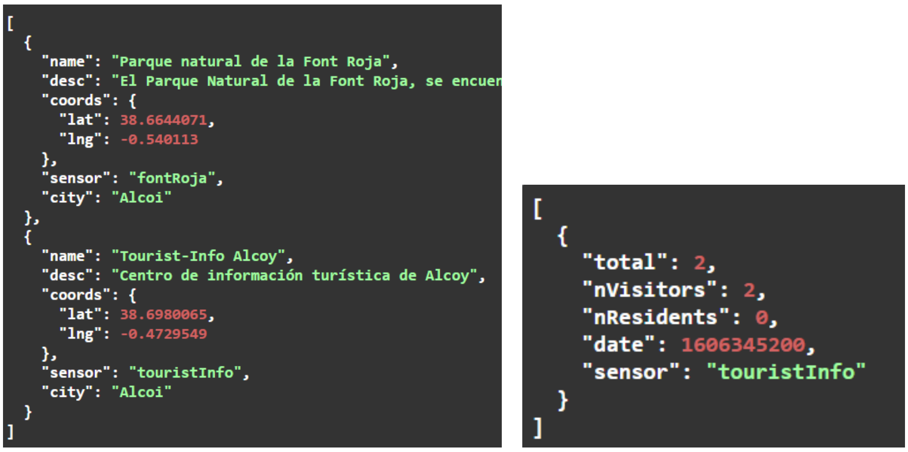
```
for each sensor in available_sensors
{
  total=0;
  visitors=0;
  residents=0;

  for each MAC detected as input_mac && MAC not stored in DB
  {
        if (isFrequentMAC(input_mac)==false)
        {
              components=getComponents(MAC);
              for each c in components
              {
                    totalSSIDs = c.length;
                    localSSIDs = c.ssidsMatch.length;
                    ratio = localSSIDs / totalSSIDs;
                    total++;

                    if (ratio > SSIDRatio
                          && totalDay (MAC) < frequentMAC)
                    {
                          residents += 1;
                    }
                    else
                    {
                          visitors += 1;
              }
        }

        insertDB (timestamp,sensor,total,residents,tourists);
  }
}
```

In the virtual machine, data is processed in order to differentiate visitors from residents. To do so, the HOPU Smart Spot devices obtain the preferred SSID data from the mobile devices within range. The HOPU Smart Spot device sends the SSID along with the date and time, as well as its identification (location) to a broker developed in Mosquitto (which implements the MQTT protocol for message management). This broker is fundamental to this system as it allows managing context information, querying and updating it (context being understood as the sensors that produce the data).

When the broker receives a message, it notifies a backend built in Node.js to process the message information and store it in the MongoDB database. In addition, this backend, at the end of each day, consolidates the collected information and stores it in MongoDB as well.

To access the consolidated daily data, an API has been implemented to provide data about the locations where the sensors are located. This API can be accessed by developers who have permission to consult its information (the permission is provided through an API key).

A Web API has been developed in order to consume data. Specifically, there are two main functions. The first one is */places* which returns the list of all the locations where sensors are placed, including a description, coordinates, and city, as well as a sensor id. Figure 4 (left) shows an excerpt of an output of the /places containing two sensors in the pilot study of Alcoi (Spain): one of them "fontRoja" located in a natural park and the other one "touristInfo" located in the Tourist Information Office both in the city of Alcoi (Spain).



```
[
  {
    "name": "Parque natural de la Font Roja",
    "desc": "El Parque Natural de la Font Roja, se encuen
    "coords": {
      "lat": 38.6644071,
      "lng": -0.540113
    },
    "sensor": "fontRoja",
    "city": "Alcoi"
  },
  {
    "name": "Tourist-Info Alcoy",
    "desc": "Centro de información turística de Alcoy",
    "coords": {
      "lat": 38.6980065,
      "lng": -0.4729549
    },
    "sensor": "touristInfo",
    "city": "Alcoi"
  }
]
```

```
[
  {
    "total": 2,
    "nVisitors": 2,
    "nResidents": 0,
    "date": 1606345200,
    "sensor": "touristInfo"
  }
]
```

Figure 4. Excerpt example for */places* (left), and for */summaryInfo* (right).

Secondly, there is the */summaryInfo* function that returns the daily report for a specific sensor (identified by its sensor id) containing the total number of people for a specific date, as well as the number of them being visitors and residents. Parameters are as follows:

- *sensor*: id of the sensor from which we want to get the data.
- *day*: date from which we want to get the data.
- If we want the data in a date range, we may indicate the date on which the range begins in the start parameter, and the date on which the range ends in the end parameter.

Figure 4 (right) shows an excerpt of an output of the */summaryInfo* containing data from a specific day obtained by the sensor "touristInfo".

3. Results

As aforementioned, a pilot study is being carried out in the city of Alcoi (Spain). It is a medium-sized city with a population of 58.994 inhabitants (2019), most of whom live in the urban area. The architecture of the urban area includes mainly 3- and 4-floor buildings.

Alcoi is a city of long industrial tradition in a process of economic diversification whereby visitor economy is playing an emergent role thanks to its rich heritage, both natural (two natural parks located in the municipal area) and cultural (a World Heritage site among other singular attractions of different historical periods). In spite of the fact of a scarce accommodation supply (585 beds) the number of visitors shows a steady growth and first signals of congestion associated with cultural events and tourist hotspots. According to this trend, the local DMO, highly committed in smart destination initiatives, is trying to develop new crowd monitoring systems, a goal reinforced by the need to guarantee social distancing in order to assure a safe visitor experience in the Covid-19 context. Specifically, Alcoi DMO needs to distinguish visitors from residents on a daily basis on several spots in the city by avoiding shortcomings of traditional burdensome surveys or other intrusive approaches such as a destination app, while privacy is preserved.

Figure 5. View of the city of Alcoi (Spain) (source: Jordi Miró).

To make a first assessment of the collection of SSIDs from the city, one vehicle and the collection system were used in a very small portion of the city for 60 minutes. The SSIDs were collected unencrypted and time-stamped.

A total of 7184 unique SSIDs were captured. To contrast the captured SSIDs with the values stored in geolocation and wardriving capture services, the database was compared with the Wigle database for the area of interest. The Figure 5 shows the Wigle map of SSIDs for the urban area of Alcoi.

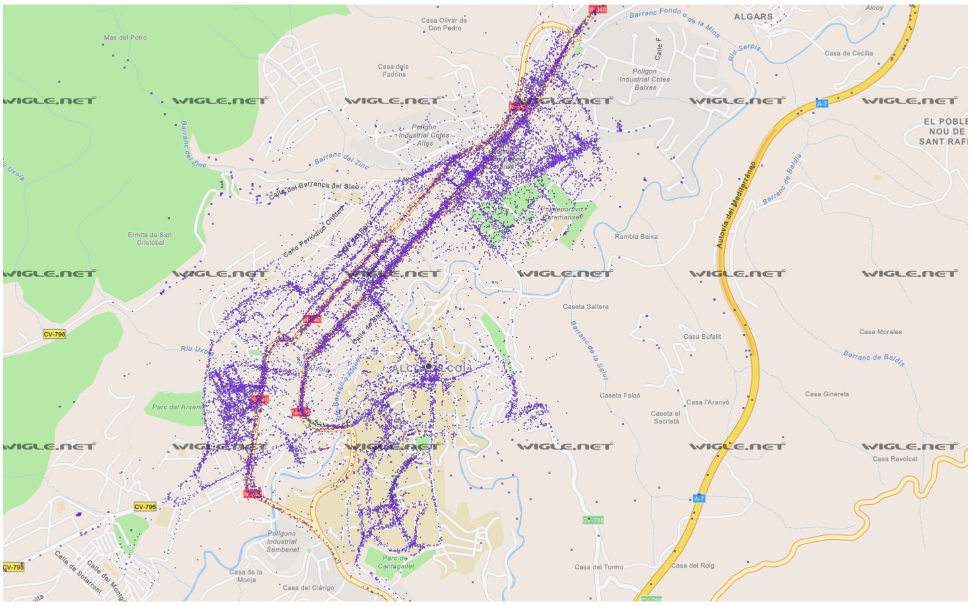


Figure 6. Wigle map of SSIDs collected in the city of Alcoi.

From the comparison with Wigle's SSIDs it is concluded that SSID names are very dynamic, changing continuously over time. This would require periodic refreshing of the local SSID database.

Another conclusion from this first analysis is that there are very common names worldwide that should be discarded in the scan, for example, the names: *AndroidAP*, *ASUS*, *dlink* or *hpsetup*. Probably, these should be discarded both for the local SSID database and HOPU Smart Spot. The database of most common SSIDs collected by Wigle would be useful to facilitate this filtering.

And another curious conclusion is that the major phone operators will repeat SSIDs of their WiFi router in many places. For example, "MOVISTAR_1DBE" (hexadecimal string modified for privacy) only allows $2^{16} = 65536$ combinations and, therefore, will be present in more than one city. In any case, we consider it valid to discriminate accepting that it compensates for the system getting some SSIDs wrong. Some router vendors have updated this issue by using longer combinations, e.g. "TP-LINK_44E4C8" (hexadecimal string modified for privacy).

Then, in order to validate our approach, it has been launched in Alcoi for 14 days (from March 1st 2021 to March 14th 2021) for 2 hours each day (from 11:00h to 13:00h). Rationale behind this timetable is that a dedicated person from the Alcoi's Tourist Info office was involved in manually conducting surveys to people that came in. Each survey collects the following information:

1. Timestamp when information was collected.
2. Town of residence
3. How many people do you travel with?
4. How many of the people you travel with have entered the Tourist Info office?
5. Do you have the WiFi on your mobile phone enabled to connect to available WiFi networks?

Table 1 shows the results of crowd monitoring using HOPU Smart Spot located in the Alcoi Tourist Info office with regard to the data collected from the surveys in the same location. It can be observed that the total number of people detected (visitors and residents) is often overestimated with our approach. However, we have to analyze specific results detecting visitors versus residents. To do so, we have developed several figures.

Table 1. Crowd monitoring results by using HOPU Smart Spots vs data collected by surveys.

	HOPU Smart Spot			**Survey**		
Date	**Visitors**	**Residents**	**Total**	**Visitors**	**Residents**	**Total**
01/03/2021	5	1	6	1	1	2
02/03/2021	1	1	2	1	1	2
03/03/2021	4	2	6	2	3	5
04/03/2021	4	0	4	3	1	4
05/03/2021	0	0	0	0	0	0
06/03/2021	11	2	13	9	1	10
07/03/2021	9	1	10	8	2	10
08/03/2021	10	4	14	1	4	5
09/03/2021	6	1	7	4	2	6
10/03/2021	7	2	9	4	2	6
11/03/2021	8	3	11	5	1	6
12/03/2021	7	1	8	6	2	8
13/03/2021	19	3	22	17	4	21
14/03/2021	2	3	5	2	0	2

Figure 7 shows the number of visitors detected by using our approach compared to the results of the surveys. Although the number of visitors is most of the time overestimated (except from 02/03, 05/03 and 14/03), the trend is shown to be equivalent. The only day the trend is broken is 08/03. Also, it can be observed that Saturdays (06/03 and 13/03) are the days in which more visitors are detected (which is an obvious result for a tourism destination).

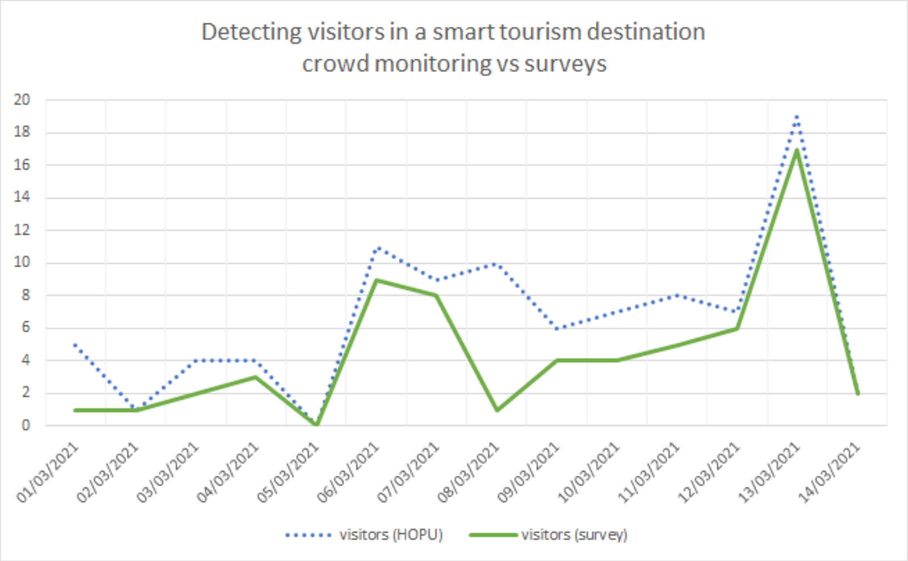


Figure 7. Using HOPU Smart Spot vs manually conducted surveys to determine number of visitors in Alcoi.

Figure 8 shows the number of residents detected by using our approach compared to the results of the surveys. The number of residents is often underestimated by our approach. Several days the number of residents detected are the same by using both mechanisms (01/03, 02/03, 05/03, 08/03, and 10/03), while on two days (06/03, and 11/03) the number of residents is overestimated by using our approach.

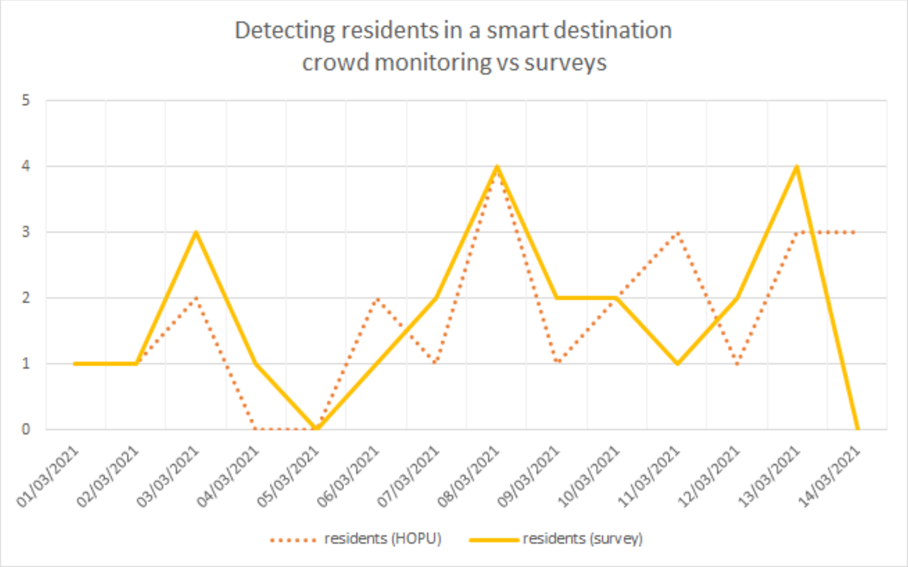


Figure 8. Using HOPU Smart Spot vs manually conducted surveys to determine number of residents in Alcoi.

Figure 9 compares the number of visitors and the number of residents detected by using both approaches (HOPU Smart Spots and surveys). It can be observed that

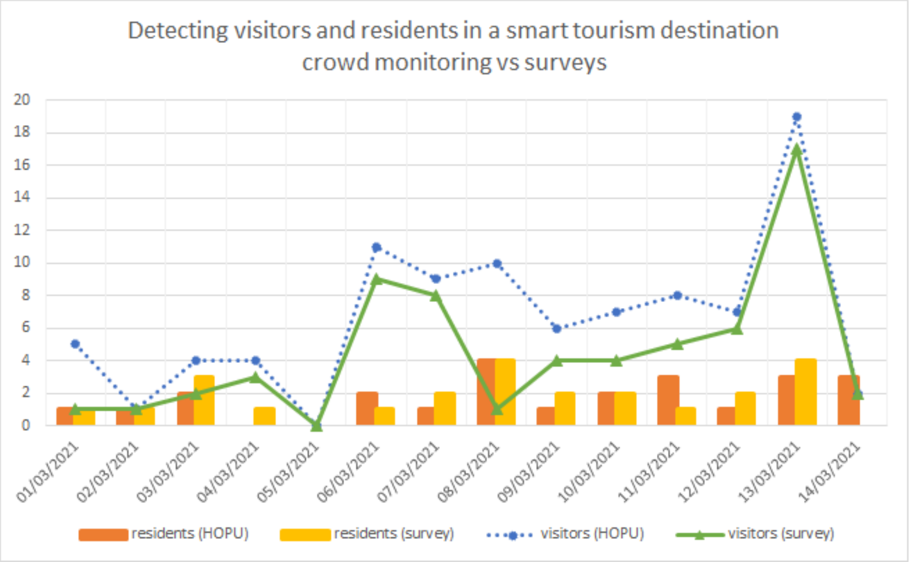


Figure 9. Comparing crowd monitoring visitors and residents in Alcoi by using HOPU Smart Spot vs manually conducted surveys.

4. Discussion

Initial results show that the processing of preferred SSIDs emitted by mobile phones is useful for distinguishing residents from visitors. According to Table 1 and Figures 7, 8 and 9, results are promising. Table 1 shows that two days (02/03, and 05/03) present correct results. In several days (specifically 01/03, 06/03, 08/03, 10/03, 11/03, and 14/03), more persons are detected with our approach based on HOPU Smart Spot than persons that actually visit the Tourist Info (according to the manually conducted survey). This is mainly caused by the WiFi scanning range of the HOPU Smart Spot. Also, several days (03/03, 04/03, 07/03, 09/03, 12/03, and 13/03) residents are detected as visitors. This is mainly caused because our list of SSIDs belonging to Alcoi's WiFi devices is incomplete. Our approach never underestimates the total number of persons detected, since we asked persons that answer the survey in the Tourist Info office whether their WiFi on their mobile phone was enabled or not. Persons with disabled WiFi connection are not considered. On the other hand, the total number of people detected is greater for our approach based on HOPU Smart Spots than for manually conducted surveys, mainly due to the need to better adjust the operating range of the device (i.e., if the operating range is too wide then mobile phones from outside the Tourist Info office could be detected).

Our approach is crowd monitoring based on WiFi scanning of mobile phones. Thus, it is a suitable tool to be used by the DMO from cities as Alcoi to prevent crowds in the pandemic times we are living in, by distinguishing visitors and residents on a daily basis on several spots in the city, while privacy is preserved. This allows Alcoi to avoid shortcomings of traditional burdensome surveys or other intrusive approaches such as a destination app. Also, it is an initial step to further consider tourist digital footprints or data traces from tourist activities, since they occur if a person can be considered a visitor as stated in [13]. Remarkably, in our approach, data analysis is done at the end of each day in order to be very privacy-conscious. Within the margins offered by the GDPR, it would be possible in the future to perform these analyses in real time to improve the responsiveness of a smart destination. Promising results of our pilot study allow us to envision other interesting future work. First, our approach should be linked to other big data sources that are already being used in smart destinations [13] (such as data from social media, booking services, destination cards and passive mobile data) in order to deploy an effective DMO. For example, in the particular case of Alcoi, techniques such as surveys in tourist offices, vehicle access to the city through number plate analysis, etc. are used and they can be considered to complement our approach. Also, extending the technique of SSID collection to other nearby cities would allow a better understanding of the flow between adjacent towns, leading to a better classification of visitors and, potentially, to a larger scale DMO tool. Finally, we are currently improving the implementation of our approach to automate tasks that require human intervention and are therefore costly in economic terms. On the one hand, it is essential that the destination takes over the collection tasks of SSIDs without incurring an additional cost; in that sense, we have decided to use the waste collection service to install our Raspberry Pi SSID collector. On the other hand, the collection of SSIDs should be as self-simulating as possible, in that sense we are developing a collector that does not need human intervention and automatically downloads the SSID data when detecting specific WiFi points of the municipal services.

Author Contributions: Conceptualization, Josep A. Ivars-Baidal, Antonio J. Jara, Jose-Norberto Mazón and Angel Perles; methodology, Josep A. Ivars-Baidal, Antonio J. Jara, Jose-Norberto Mazón and Angel Perles; hardware and software, Alberto Berenguer, David Fernández, Andrea Gómez-Oliva, and Jaime Laborda; pilot study, Alberto Berenguer, David Fernández, Andrea Gómez-Oliva, Josep A. Ivars-Baidal, and Antonio J. Jara; validation, Alberto Berenguer, Jose-Norberto Mazón and Angel Perles; writing—original draft preparation, Alberto Berenguer, David Fernández, Andrea Gómez-Oliva, Josep A. Ivars-Baidal, Antonio J. Jara, Jaime Laborda, Jose-Norberto Mazón, and Angel Perles. All authors have read and agreed to the published version of the manuscript. Authors are alphabetically listed in the paper.

Funding: This research has been carried out within the framework of the research Project Alcoi Tourist Lab, cofunded by the Alcoi City Council & the Valencian Innovation Agency. This research was also partially funded by project UAPOSTCOVID19-10 from University of Alicante.

Acknowledgments: We would like to thank Paola Pons and Pedro Ramiro from the Alcoi City Council involved in the Project Alcoi Tourist Lab.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gretzel, U.; Sigala, M.; Xiang, Z.; Koo, C. Smart tourism: foundations and developments. *Electronic Markets* **2015** 25(3), 179–188. https://doi.org/10.1007/s12525-015-0196-8
2. Gretzel, U.; Werthner, H.; Koo, C.; Lamsfus, C. Conceptual foundations for understanding smart tourism ecosystems. *Computers in Human Behavior* **2015** 50, 558–563. https://doi.org/10.1016/j.chb.2015.03.043
3. Ivars-Baidal, J. A.; Celdrán-Bernabeu, M. A.; Mazón, J.-N.; Perles-Ivars, A. F. Smart destinations and the evolution of ICTs: a new scenario for destination management? *Current Issues in Tourism*. **2020**, 22(13), 1581-1600. https://doi.org/10.1080/13683500.2017.1388771
4. Boes, K.; Buhalis, D.; Inversini, A. Smart tourism destinations: ecosystems for tourism destination competitiveness. *International Journal of Tourism Cities* **2016** 2(2), 108–124. https://doi.org/10.1108/IJTC-12-2015-0032
5. Buhalis, D.; Amaranggana, A. Smart Tourism Destinations. In Information and Communication Technologies in Tourism 2014. Proceedings of the International Conference, Dublin; January 21-24, 2014; Z. Xiang & I. Tussyadiah Eds; Publisher: Springer International Publishing, 553–564. Retrieved from http://link.springer.com/10.1007/978-3-319-03973-2_40
6. Hardy, A. *Tracking Tourists. Movement and Mobility*; Publisher: Goodfellow Publishers, UK, 2020.

7. García-Hernández, M.; Ivars-Baidal, J.; Mendoza de Miguel, S. Overtourism in urban destinations: the myth of smart solutions. *Boletín de la Asociación de Geógrafos Españoles* **2019** 83(2830), 1–38. http://dx.doi.org/10.21138/bage.2830
8. Yang, T.; Yan, Z.; Wen, J. Impact of COVID-19 Pandemic on Smart Tourism. Advances in Economics, Business and Management Research. Proceedings of the 5th Asia-Pacific Conference on Economic Research and Management Innovation (ERMI 2021), on-line, January 31, 2021, 2021. Retrieved from https://www.atlantis-press.com/proceedings/ermi-21/125952872
9. 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/EC (General Data Protection Regulation). The European Parliament and The Council of the European Union. http://data.europa.eu/eli/reg/2016/679/oj (accessed on 27 March 2021)
10. Singh, U.; Determe, J. F.; Horlin, F.; De Doncker, P. Crowd Monitoring: State-of-the-Art and Future Directions. *IETE Technical Review* **2020**, 1-17 https://doi.org/10.1080/02564602.2020.1803152
11. ESP32-Paxcounter. Available online: https://github.com/cyberman54/ESP32-Paxcounter (accessed on 27 March 2021)
12. Oliveira, L.; Schneider, D.; De Souza, J.; Shen, W. Mobile device detection through WiFi probe request analysis. *IEEE Access* **2019** 7 98579-98588. https://doi.org/10.1109/ACCESS.2019.2925406
13. Reif, J.; Schmücker, D. Exploring new ways of visitor tracking using big data sources: Opportunities and limits of passive mobile data for tourism. *Journal of Destination Marketing & Management* **2020** 18, 100481. https://doi.org/10.1016/j.jdmm.2020.100481
14. HOPU Smart Spots. HOP Ubiquitous S.L. https://smartcities.hopu.eu/smart-spot.html (accessed on 27 March 2021)
15. Wi-Fi positioning system. Wikipedia. https://en.m.wikipedia.org/wiki/Wi-Fi_positioning_system (accessed on 27 March 2021)
16. Lischka, K. Google-Debatte: Datenschützer kritisieren W-Lan-Kartografie. *Der Spiegel on-line*. 2021. Retrieved from http://www.spiegel.de/netzwelt/web/0,1518,690600,00.html
17. WiGLE: Wireless Network Mapping. https://wigle.net/ (accessed on 27 March 2021)