A Survey of Techniques for Automatically Sensing the Behavior of a Crowd

ADRIANA DRAGHICI, University Politehnica Bucharest
MAARTEN VAN STEEN, University of Twente

Crowd-centric research is receiving increasingly more attention as datasets on crowd behavior are becoming readily available. We have come to a point where many of the models on pedestrian analytics introduced in the last decade, which have mostly not been validated, can now be tested using real-world datasets. In this survey, we concentrate exclusively on automatically gathering such datasets, which we refer to as sensing the behavior of pedestrians. We roughly distinguish two approaches: one that requires users to explicitly use local applications and wearables, and one that scans the presence of handheld devices such as smartphones. We come to the conclusion that despite the numerous reports in popular media, relatively few groups have been looking into practical solutions for sensing pedestrian behavior. Moreover, we find that much work is still needed, in particular when it comes to combining privacy, transparency, scalability, and ease of deployment. We report on over 90 relevant articles and discuss and compare in detail 30 reports on sensing pedestrian behavior.

CCS Concepts: • **General and reference** → **Surveys and overviews**; • **Information systems** → **Spatial-temporal systems**; • **Human-centered computing** → **Ubiquitous and mobile devices**; • **Computer systems organization** → *Sensors and actuators*;

Additional Key Words and Phrases: Pedestrian sensing, pedestrian tracking, crowd sensing

ACM Reference format:
Adriana Draghici and Maarten van Steen. 2018. A Survey of Techniques for Automatically Sensing the Behavior of a Crowd. *ACM Comput. Surv.* 51, 1, Article 21 (February 2018), 40 pages.
https://doi.org/10.1145/3129343

1 INTRODUCTION

Crowd-centric research has been around for more than a decade and has gradually become an established interdisciplinary field of its own. With a multitude of stakeholders, a wide range of applicable scenarios, and many different problems and approaches toward solutions, it has also become a complex field of research.

For example, crowd-centric research covers indoor and outdoor pedestrian tracking and ranges from small buildings to large shopping malls to huge festivals. A wealth of models have been developed for purposes of merely understanding crowd behavior, realistically simulating such behavior for visualization purposes, or actually predicting future behavior. There are myriad reasons for

Authors' addresses: A. Draghici, University Politehnica Bucharest, Faculty of Automatic Control and System Engineering, 313 Splaiul Independentei, 060042, Bucharest, Romania; email: adriana.draghici@cs.pub.ro; M. van Steen, University of Twente, Institute on the Digital Society, Drienerlolaan 5, 7522 NB Enschede, The Netherlands; email: m.r.vansteen@utwente.nl.

https://doi.org/10.1145/3129343

wanting to understand or predict pedestrian behavior: safety, marketing, planning, and general management to name just a few.

In this era of data-driven research, there is an increasing trend toward developing crowd-behavior models using real-world data. Unfortunately, as concluded from a recent extensive literature survey [81], data-driven research for modeling crowd behavior is by far common practice. This lack of research can be explained by the difficulty of obtaining datasets, especially for very large crowds. Furthermore, the quality of available datasets is often unclear as cleaning and sanitizing raw data has its own problems [12]. Yet, the need for high-quality datasets capturing the behavior of crowds is undisputed.

In this article, we investigate the various methods and techniques for capturing crowd behavior through physical sensors that record spatiotemporal features such as densities and movements. We exclusively focus on alternatives to CCTV and other video-based techniques; in particular, we consider radio-based infrastructures such as WiFi-tracking systems and systems using Bluetooth beacons. Our goal is to provide an overview of ways to automatically *sense* the behavior of a crowd. In particular, we focus on automatically detecting information on positioning, tracking, and measuring collections of people. This is what we refer to as *sensing crowd behavior*. This sensing is not to be confused with *crowdsensing*, which is a form of urban crowdsourcing, a method of using a person's phone as a sensing node that gathers data about surrounding phenomena [25]. Throughout this survey, crowd sensing always refers to sensing a crowd, unless stated otherwise.

Until recently, many sensing solutions relied on custom nodes or networks of devices. The current trend is to leverage the sensing capabilities of wearable devices and notably smartphones using participatory applications. For example, it is now relatively easy to detect the presence of nearby devices, obtain movement or location data, or acquire all sorts of local environmental data. Combining such data with information from social media turns a smartphone into an extremely powerful and versatile multisensing device.

We distinguish three different categories for using a wearable multisensing device. First, in the case of human-centric (also called people-centric) sensing, the goal is to collect data on personal traits: movement, activity, stress, and so on. Second, with environment-centric sensing, the goal is to capture information on the surroundings of a person, such as data on weather, pollution, traffic, and so forth. Finally, the third category involves crowd-centric sensing, which emphasizes collecting spatiotemporal data on the behavior of groups of people typically aiming at estimating the size of a crowd, local densities, flows, and so on. In this article, we focus on crowd-centric sensing.

Admittedly, the boundaries between these categories are not always clear, in particular when considering that in many cases the same sensors are used. Nevertheless, when concentrating on the *purpose* of sensing, distinctions arise. For one, in the case of crowd-centric sensing, it is not an individual person who is generally the object of study, but rather the crowd as a whole. As a result, there is generally more emphasis on gathering aggregated statistics, and the tolerance for having to deal with noisy data is much higher than, for example, with human-centric sensing. Likewise, where scalability is an inherent design issue for crowd-centric sensing, this is generally much less the case for environment-centric or human-centric sensing. Scalability can easily lead to radically different designs if one is targeting the behavior of millions of people. Thus, while some of the challenges we identify in this survey have a common ground with other sensing domains, the fact that they are targeted to capturing the behavior of crowds raises many new interesting research questions.

We identify two types of systems for crowd-centric sensing: application-driven and infrastructure-based systems. Application-driven systems essentially make use of wearable devices for sensing the behavior of a crowd. A typical example is using smartphones to collect data on the number and location of neighboring devices. Infrastructure-based systems typically use statically placed sensors that scan for wearable devices (and no more than that). A well-known example is the use of WiFi scanners for detecting the presence and recurrence of WiFi-enabled smartphones. Of course, hybrid forms exist as well. Both types of systems can be either participatory or opportunistic and can be applied to several types of indoor and outdoor environments. Our survey focuses on the whole spectrum of solutions and identifies their architectural approaches and challenges.

Social media traces, collected from specific platforms (e.g., Foursquare) or using dedicated applications, can also provide crowd-related data. This is a different approach than the one we are focusing on and warrants a separate survey centered more on data analysis. We concentrate only on minimal-intrusion sensors for detecting the physical presence of devices and do not dwell on the semantics of social media. Nonetheless, we included application-driven sensing systems that analyzed social media data in addition to the dataset collected using the mobile devices' sensors because they used it to validate their field experiments.

The sensing modalities employed by the systems we surveyed are also used for localization and tracking of individuals. Although we also mention notable papers on these topics, our target is the systems that collect spatiotemporal datasets that can describe crowds. The papers that just analyze crowd data without describing the sensing part (technologies, experiments, methods) are not the focus of this article.

We reviewed 93 papers on topics related to sensing crowds, falling into the categories described below. Most of them present sensing systems that collect and analyze mobility data. Although they rely on field experiments using mobile applications or deployed sensors, none consists of an operational system used on a daily basis. The sensing solutions that were operational a few years ago such as the mobile applications CitySense [44], VibN [50], and CoenoSense [83] are no longer available on current mobile platforms. We distinguish the following *types of papers*:

- Papers on urban sensing systems, such as pedestrian monitoring using applications or sensing infrastructures. In most cases, analysis focuses on pedestrian flow throughout the city and on determining popular places.
- Papers on indoor sensing systems. These mostly concern infrastructure-based systems for tracking people inside buildings. The data can be used for analyzing flows, patterns, and densities, but usually the authors focus on only one type of pattern. They also present pre-experiment tests and calibrations.
- Papers on event monitoring, both indoor and outdoor, and at varying scales. These type of papers focus both on the experiment and on the analysis of the collected data.
- Papers on frameworks for participatory sensing applications.
- Position papers on sensing architectures and related topics such as privacy, evaluation methodologies, and heterogeneity of sources.

Less than half of the urban and indoor-sensing and event-monitoring papers are completely focusing on sensing mechanisms for crowds, a subset we will refer to as **spot-on papers**. These present real-life deployments and their subsequent analyses. They provide details on the sensing technologies, methodologies, and implementation, thus representing the main focus of our survey. These systems are subject to various challenges and tradeoffs particular to sensing crowds. We have classified them based on how they address these issues. This classification performed in Section 5 covers the main architectural and nonarchitectural criteria for acquiring crowd mobility data: security and privacy, ease of deployment, scalability, incentives, transparency, and resource consumption. Accuracy is another criterion we considered, but it is difficult to quantify in a rating due to the variety of analysis methods and metrics encountered in the surveyed papers.

We reviewed and classified papers related to crowd sensing following a survey methodology that consisted of five phases: paper selection, general characteristics classification, crowd-sensing characteristics classification, spot-on systems identification, and comparative evaluation of all representative papers. The differences between the second and the third phase consist of the type of information we extracted from the papers. In the second phase, we identified characteristics such as technologies, experiments, purpose, and beneficiaries. In the third phase, we proposed seven main features for crowd-sensing systems and evaluation criteria for the sensing architectures.

This methodology influenced the organization of the article. In the following two sections, we apply our classification criteria on the applications and infrastructures presented in the set of papers we selected. In Section 5, we discuss the most representative papers and compare them based on the features presented in Section 2. We conclude in Section 6 by discussing our view on the current state of crowd sensing and the trends and challenges we noticed in the papers we surveyed. Further information can be found in Draghici [19].

2 KEY FEATURES OF CROWD-SENSING SYSTEMS

In this survey, we focus on systems that sense the behavior of crowds, in particular those systems that are an alternative to video-based solutions. We focus on the particularities of crowd-centric sensing solutions and their similarities and differences with traditional sensing systems. We also identify the main properties that should be taken into account when designing a system for sensing the crowd. In Section 5, we discuss such existing systems from the perspective of these properties.

2.1 Architectural Considerations

Both application-driven and infrastructure-based systems are relying on a centralized architecture with devices performing the sensing (or some of the processing) and transmitting the data to a server for storage, analysis, and presentation. For both types of systems, coping with heterogeneity is important. In application-driven systems, the sensing devices are the main source of heterogeneity: different platforms have different sensing APIs and restricting policies, but also different sensing, processing, and communication hardware. Blunck et al. [10] also consider the users as a source of heterogeneity due to demographics and variations in application and device usage. For the infrastructure-based systems, heterogeneity comes mainly from the sensed devices, such as differences in signal strength or scanning periods.

Zooming in on the architecture, we encounter several processing, storage, and communication models. Processing is performed either locally on the device, remote on the server, or both. For application-driven systems, the policies dictating this choice are generally driven by energy consumption requirements. How the collected data is stored depends on the storage capabilities of the device but also on the privacy policies of the application. While the processing model is fixed, the storage model can generally be customized by the user. We encounter these models also in other sensing systems, but there are subtle differences. For instance, in a participatory sensing application for fitness, the user may opt to store the data only locally and never transmit it to a server for further processing. This is obviously not an option in the face of building a global view on crowd behavior.

Awareness of energy and resource consumption also influences the communication mode and sensing strategies. Sensing can either be performed continuously in the background or triggered by an input from the user. Energy-aware applications adjust the sampling rate or even the sensors used in order to reduce the consumption.

Typically, sensing devices are assumed to always have Internet connectivity and to almost instantly transmit data to the server. When continuous connectivity cannot be guaranteed, data is gathered after an event from local storage, as in [70]. A drawback is that participants can then not receive real-time feedback on global crowd behavior.

Table 1. Common Crowd-sensing Technologies from a Sample of 67 Systems (27 Application-driven, 40 Infrastructure-based), Together with the Number (and Fraction) of Surveyed Solutions Employing them. Some of the Systems Use Multiple Technologies

Type	Bluetooth	WiFi	GPS	Microphone	Camera	Motion Sensors
Application	10 (0.37)	5 (0.19)	16 (0.59)	7 (0.25)	3 (0.11)	11 (0.41)
Infrastructure	17 (0.42)	21 (0.53)	3 (0.07)	0 (0.00)	0 (0.00)	8 (0.20)

Mobile applications for sensing the crowd present more diverse communication strategies than the infrastructure-based systems. They are usually closely connected with the sensing model and can be triggered either by the device, by the server, or in some cases even by another device. A device may wait for tasks from the server, start the data collection, and return the results, or may simply publish data, without a specific request, whenever a WiFi connection is available.

2.2 Sensing Modalities

The sensing literature offers comprehensive surveys [25, 45, 73] on the technologies used for acquiring data on human and environment traits. The systems for sensing the crowd leverage some of these technologies to obtain data about the presence, the count, and the movement of people. We identified several sensing modalities and their corresponding technologies. Table 1 presents the technologies behind these modalities and the number of surveyed solutions for each of them.

Choosing the right modality is important, if only for reasons of energy consumption, costs of resources, data granularity, and implementation and deployment restrictions. Some, such as energy impact and implementation restrictions, are more relevant to application-driven solutions. The energy consumption is dependent on the type of sensors, the hardware platform and the operating system, the API of the mobile device, and the collection method.

Some crowd-sensing solutions based on participatory applications enhance their analysis by combining data from several modalities with social media information. CrowdSense@Place [14] crowdsources the gathering of data about the urban environment, and while it is not strictly a system for sensing the crowds, with a significant user base it can provide information on crowd densities and movement patterns. In Chon et al. [14], this system was used in an experiment with just 85 participants, yet they managed to gather data about visit counts and app usage. Note that this kind of information cannot offer any global indication about a crowd. This hybrid approach can be applied to sensing crowds especially in the case of city-scale events or for determining popular places, but we have not yet encountered crowd-sensing frameworks and applications that support it.

2.3 Maturity of Crowd-Centric Sensing Solutions

Crowd-centric sensing handles large numbers of participants, heterogeneous devices, and various types of environments, imposing challenges on the testing and evaluation processes. Some of the sensing technologies described above have been employed, tested, and optimized on tracking and localization of individuals. For handling crowds, the collected data must be representative and valid for more than just an individual. The systems for sensing a crowd relying on outdoor experiments outnumber the ones analyzing datasets collected through small-scale lab experiments and simulations. In the papers we surveyed, the testing and evaluation mostly depended on the purpose of the presented solution. Some were built just as a basis for a particular type of analysis, and some for demonstrating the feasibility of a particular technology or for comparing technologies (such as by Abedi et al. [1] and Schauer et al. [66], who compare Bluetooth and WiFi). Others have been developed for monitoring for only a certain amount of time, such as during specific events of various scales, from indoor exhibitions or conferences to city-scale festivals.

Usually, crowd-centric solutions consist of the following stages: pre-experiment calibration, deployment (i.e., actual sensing), and finally data analysis. Most papers do not address the first stage, with a few exceptions in case of infrastructure-based systems using radio-based modalities.

The deployment stage consists of one or more field experiments, either instrumented or not. In the former case, the experiment consists of the monitoring of a few volunteers (usually less than 20) equipped with phones or other sensing devices, sometimes following a specific script. In noninstrumented cases, either an application is made available to any user or sensing devices are deployed to monitor any person passing by. While the latter deployment usually produces the largest datasets, these datasets are also more problematic to analyze and validate. Moreover, such experiments, especially those covering a large area or with a large number of users (typically over 1,000), are more prone to data quality problems and unexpected events.

One of the challenging parts of the validation process is collecting ground-truth data necessary for evaluating the accuracy of the experiment. For instrumented approaches with a few dozen participants, it is relatively easy to determine the ground truth, even by using human observers. For more complex experiments, ground-truth data is collected either by video monitoring (e.g., [38, 83]), manual observations [14, 28, 32, 46, 55, 57, 59], additional sensing modalities such as GPS [53, 80], motion detectors [24], location-specific modalities (such as turnstiles [20] or boarding-pass scans [66]), or social media check-ins [13, 14]. Almost half of what we termed **spot-on solutions** on sensing crowds do not even present a ground-truth strategy, comparing their results with various statistics (e.g., known distributions on cell phone usage) or identifying relevant patterns (rush hours, diurnal patterns).

Despite the staging costs (devices, rewards for participants), the instrumented experiments seem to be the common method for demonstrating the feasibility of a certain crowd analysis method or the collection accuracy of a certain sensing modality. The question remains, though, whether these sensing mechanisms scale. For infrastructure-based systems, we have the problem of coverage and deployment costs. For application-driven approaches, we have nontechnical challenges such as attracting users, or additional technical challenges regarding privacy, security, and resource consumption.

2.4 Features for Evaluation

Security and privacy. The idea of a system that continuously collects data on pedestrians raises ethical, privacy, and security concerns. Threats can be both internal and external and can target the sensing, the data collection (task communication and results reporting), the local and remote storage, and even the presentation (e.g., when querying for statistics of currently "hot" places). The *Privacy* criterion in our classification encompasses anonymization, security, and access and sharing policies.

In participatory sensing applications, privacy guarantees that users have control over their data. Their collected and inferred information is protected and not available to other users or parties. For such applications, anonymization is not always a requirement, especially for localization and tracking applications, but is preferable in case of collecting data on crowds.

In infrastructure-based crowd-sensing systems, people have much less control over their participation. In this case, anonymization is often a requirement and consists of stripping the datasets of context and demographic information. Some of the systems we reviewed used address hashing, a technique that is possible to de-anonymize unless it is coupled with other privacy-preserving schemes [13].

Incentives. Sensing the behavior of a crowd generally requires participation of many people. When this participation has to be solicited, incentives become important.

The incentives mechanisms for application-driven systems are the ones usually employed in participatory systems. Restuccia et al. [62] provide a recent survey and Lee [41] an in-depth study of the economic models. Arakawa and Matsuda [3] present a study of gamification mechanisms for urban participatory sensing as an alternative to monetary incentives. Crowd-centric application-driven systems usually rely on non-auction-based mechanisms and provide monetary incentives or application-specific ones, which include gamification, integration with social media, access to certain content, or analysis results (e.g., the user sees how crowded a specific place is only if he or she agrees to share his or her location). The incentives, while closely coupled with the privacy concerns, are also important when talking about the deployment or how the application is made available to the users. Embedding solutions into an existing app [9] can make a huge difference in comparison to a separate app [71].

Ease of deployment. We also consider the way the system is deployed and its maintenance requirements, distribution, and marketing efforts. The sensing systems we reviewed presented very briefly the server-side deployment or costs, with the deployment discussions focusing on the sensing devices. Mobile-driven solutions need to make the application available through official channels, such as Google Play on Android and rent server resources in the cloud. The amount of effort shifts from the deployment to the implementation and maintenance side. For infrastructure-based systems, the deployment is more costly since most cases require custom sensing devices covering a large area but need less marketing and implementation effort and, if properly placed, can produce large datasets immediately, while application-driven systems require time to build the user base.

Scalability. We consider a sensing system scalable if it can be easily adapted and without significant costs to support larger areas, more users, and extended periods of time. This aspect considers the impact that scaling has both on sensing infrastructure costs and on processing and storage resources. Some of the sensing systems we analyzed were also designed for a small number of participants or low densities, the analysis becoming less accurate when this number increased. Also, for mobile-driven systems, the analysis and filtering need to account for similar reporting from persons close by. The stress on the server-side systems due to an increase of data that needs to be received, stored, and processed is not discussed in the reviewed sensing solutions.

This topic is addressed in a few papers only. For example, Kannan et al. [33] include a formal discussion on the scalability of the tone-based crowd-counting system they propose. They also discuss the ease of deployment and energy efficiency criteria.

Transparency. What is the level of awareness of the user about the sensing campaign and data collection? Sensing infrastructures that just monitor passing-by devices are considered to be almost entirely transparent to the users, in contrast to mobile applications that constantly require interaction with the user. Transparency is particularly challenging in application-driven systems in which usability comes into play while ensuring a minimal effect on other applications and resources. Transparency is also affected by the sensing modalities used in the smartphone app. Due to security reasons, the mobile platforms' APIs impose restrictions on accessing and enabling these modalities, which affect the transparency by requesting user input.

Resource consumption. A serious research challenge in many sensing systems, and also in those for sensing crowds, is controlling resource usage. This holds not only for devices but also for server-side resources, being closely connected to scalability, transparency, and accuracy. Application-driven systems usually tackle energy efficiency by implementing policies for minimizing the consumption, for example, dynamically adapting the rate for acquiring the location based on the user's movements [7, 30].

Related is the system's complexity: a good application that needs resources for collecting fine-grained mobility data, provides incentives, and presents results is preferred to a simple application that collects less accurate datasets and does little to attract the users.

Accuracy. For systems on localization and tracking of individuals, positioning accuracy is a main concern. On the other hand, in crowd-centric systems, we see a large spectrum of characteristics considered by their researchers and developers (as discussed in Section 4.3), and the metrics are more varied. Most of the spot-on papers we surveyed presented their analysis results but in various degrees, some just presenting counts or simple statistics about the device vendors. This criterion encompasses the types of analysis, the filtering needed to clean up the data, the metrics, and (if any) the validation mechanism. In addition to the evaluation results, we consider whether or not their choice of technology and deployment is capable of providing representative datasets. For instance, we have seen significant changes for radio-based modalities due to rapid changes in mobile platforms. We discuss accuracy throughout the next sections applied to the systems we surveyed, but due to its variance, we do not employ a rating system as for the rest of criteria.

3 APPLICATION-DRIVEN SENSING

3.1 Architecture

In this section, we first introduce a complete architecture for application-driven crowd-centric systems, which encompasses building blocks for both the device and back end, as illustrated in Figure 2 in the appendix. We then provide examples of existing systems that successfully implemented similar architectures.

3.1.1 Components of the Client. A mobile application generally consists of background components and UI components. A few solutions provided only the background components, as services that can be used by various applications. Decoupled and modular architectures are more versatile and can be integrated with multiple applications. For instance, a service that provides sensing and communication can be used by applications designed for different events or festivals, or various applications created for the same event [71].

Sensing. Most important is the sensing component, since it is the one collecting the raw data from the sensors or communication interfaces (for proximity detection). The application can additionally include energy-aware policies such as dynamically adjusted sampling rates based on the movement type or context, or merely sampling on demand. Applications using multiple sensing modalities can also perform sensor selection in order to alternate the sensors having a low energy cost with the high-power radio or location providers, possibly trading energy for accuracy.

Processing. The processing component is optional, some solutions preferring to do the processing only at the server in order to have a lower impact on the device's resources. Others perform some basic filtering and anonymization of the data before sending it. We also encountered systems performing a significant amount of processing on the device, for instance, Miluzzo et al. [49], which executes audio classification and activity recognition based on motion sensors. Their approach is driven by privacy considerations, the raw data being stored only temporarily while processed and the server receiving only the results of the processing for further analysis and integration with other data streams.

Privacy enforcement. The optional privacy enforcement component generally consists of a series of mechanisms for implementing privacy policies during the sensing, processing, or communication. Typical examples include constraints on the area in which sensing is active or on how a certain modality is used. When applications use several sensing modalities, they often implement an on-demand policy for obvious privacy-sensitive modalities (such as camera and microphone) and a continuous collection policy for motion sensors or location providers.

Data storage policies are usually driven by the privacy settings of the system, settings that are either established by the application logic or configurable by the user. Storage policies are often coupled with processing, especially in applications that collect audio streams, filter them of any identification content, and then remove the raw samples, storing just the processed data.

The communication component may also strip the reported data of identification features, the most common procedure being to hash the addresses involved or to not send details about the user and its device. How effective such policies are is questionable (see, e.g., Vanhoef et al. [74]). This anonymization step, employed by most infrastructure-based systems for sensing the crowds, is not that often encountered in participatory applications. The fact that they use social media integration as an incentive makes their users share their identity with the back-end services. In these cases, the server needs to protect the stored data and to guarantee not sharing the information to third parties without the user's authorization. In fact, some studies suggest that users are not that concerned with sharing their location history or other sensor data when they are in public places [8, 14, 49].

Communication. The communication component is responsible for reporting data to the server and receiving tasks or other information related to the collection campaign. Depending on the implementation, the sensing component may use some of this component's functionalities, for example, when it needs to use communication interfaces to detect neighboring devices. Likewise, some systems use short-range communication not just for detection but also for enabling collaboration between devices.

There are a few infrastructure-based systems that lack a communication component, saving collected data on local storage in order to be accessed only after the event [70]. Also, several participatory applications designed for sensing personal traits may not include a communication component. However, the nature of the crowd-related data requires collecting and aggregating samples from multiple users and locations. Even when the system is completely decentralized and the mobile device collects data in an ad hoc manner from the devices it encounters, it will eventually need to communicate its findings to a logically centralized service.

Presentation and user-controlled settings. The design of the user interface is mostly driven by transparency, usability, and incentive requirements. Interestingly, most applications do not provide information on current crowd conditions and instead focus more on gathering input (including gamification) and provide only general event information [9, 71, 83].

Even when users do not have access to the sensing campaign results, they must be informed of the collection for reasons of imposed resource usage and invasiveness on privacy. A user should have the option to opt out entirely of the collection process and be offered support for configuring issues like sampling rate, turning on and off the sensing, deciding how long the data is locally stored, or if the communication is performed only when connected to open networks, to name a few. Most applications in our survey offer such capabilities.

Incentives. Incentivizing users remains a challenging area, notably in participatory systems [62]. For sensing crowds, incentivizing tactics generally encourage participation in data collection by engaging the users either with application-specific features or with gamification mechanisms. Out of the app-driven solutions we have surveyed, just one application had incentives as a primary design feature [8, 9], offering a virtual-trophy-collecting game. The authors also show a high interest in studying incentivizing mechanisms and even surveyed the users about the gamification elements they included.

A few applications used incentives as a means just to reward volunteers in a field experiment. Monetary incentives are a viable option as well, but we have not seen them be integrated into real deployments of mobile crowd-sensing applications.

External services and applications. Many systems for crowd sensing can be integrated with other services or applications for presentation purposes, storage, sharing, or authentication. For example, applications that offer real-time information on crowd densities are often linked to the Google APIs for map integration and location awareness. The application may also offer options for synchronizing data with services such as Dropbox or to share information via social media.

3.2 Components at the Server

Crowd management solutions can be logically split into four major subsystems [81]: sensing, mining, prediction, and intervention selection. Many of the solutions that we have included in our survey also address elements of subsystems other than the one for sensing. However, in this article, we confine ourselves exclusively to the sensing subsystem. In this section, we zoom into this subsystem's organization at the server side.

Communication. The communication component is primarily responsible for asynchronously receiving data from the devices. Depending on the design tactics, the server may send requests (tasks) for triggering data collection or for obtaining collected data. It can also answer requests for processed or aggregated data. This is the case with applications that provide information on crowd conditions (e.g., the densities in a given area during the last week) or use the user's server-side stored data in their local processing (e.g., for pedestrian dead-reckoning techniques).

Privacy enforcement. Privacy policies can be enforced at both the client side and the server side. For crowd-centric sensing, we generally do not need the identities of participants. At the client side, data can be stripped of identification before being sent to the server. Otherwise, hashing methods can be applied on the server. When the system is designed to know a user's identity, it can enforce access policies for the user's data. When querying for crowd conditions, the client receives just aggregates (e.g., visit counts in a certain area in a given timeframe) and never information on specific people.

Clearly, to what extent privacy enforcement at the server is effective remains an open question, certainly in light of potential security attacks. None of the surveyed sensing systems had by far an adequate solution.

Control, storage, and processing. The control component is the one responsible for the system's logic tier. It sends tasks to the application through the communication component, it interprets the requests from the application, and it controls the processing stages: filtering, data mining, and visualizations. In general, it forms the core of the crowd management system.

Presentation and external services. The system can also offer a presentation component, which provides statistics and visualizations of the collected data via a web interface. These can be publicly available or just private to the users, crowd operators, and developers. Similar to the mobile application, the presentation component can be integrated with external services for maps, location information, or even graph plotting tools.

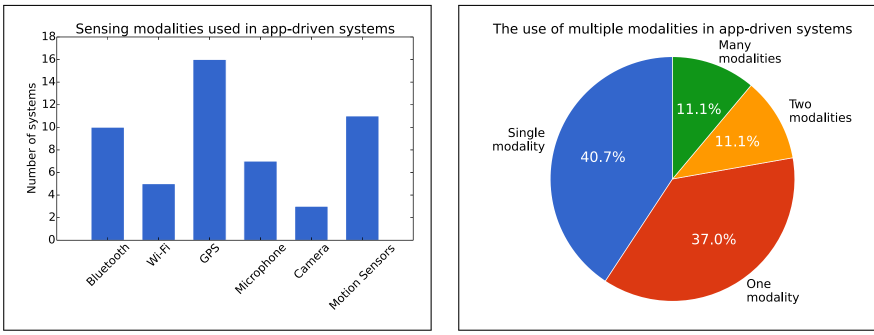


Fig. 1. Sensing modalities.

3.3 Sensing Modalities

The applications designed for sensing crowd characteristics use mostly one or two sensing modalities, with location providers being the easiest and straightforward option. As seen in Figure 1, out of the 27 application-driven systems we have surveyed, most of them use GPS or motion sensors. For energy considerations, some combine the location acquisition with the data from motion sensors (mostly accelerometer and compass) in order to dynamically adjust the location provider's collection rate. The strategies for adjusting the sampling rate consider the user's speed (type of movement), traveled distance, and heading.

While in theory these strategies should work, the implementation of such policies needs to adapt to the restrictions of the current mobile platforms. The mobile market is extremely dynamic and heterogeneous, and the available APIs constantly add more restrictive policies to protect privacy or reduce energy consumption. One such restriction is available on Android, where the sensor data can be continuously collected but is not transmitted to any server when the screen is off. This is an impediment for the applications that need to acquire the location after a certain number of steps or traveled distance. The four systems that considered such strategies [7, 30, 37, 50] either used an older, less restrictive platform [37, 49] or implemented a prototype used in a small-scale instrumented experiment. Höpfner and Schirmer [30] propose several workarounds, even one based on static movement profiles, and show promising results in the evaluation against the SDK's default policy. Their implementation requires almost half of the number of SDK requests but has lower positioning accuracy (e.g., 11m instead of 4.5m or 7.5m). The latest version of EnTracked [7] has a more in-depth analysis of this tradeoff between reducing the energy consumption and giving up some of the positioning accuracy.

3.4 Frameworks

In our study, we have encountered mostly frameworks that are only indirectly linked to sensing crowd behavior. These frameworks are built primarily for participatory sensing. They address energy efficiency, privacy, and participant recruitment. We have also encountered papers offering frameworks and at a conceptual level [26, 29], or valuable insight on architectural tactics for mobile sensing [36].

One of the most relevant examples for our study is Medusa [60], which allows developers to define tasks that can be used for sensing characteristics of a crowd. This framework supports all the sensing modalities we have presented in Section 2 and its authors also discuss place-centric applications using them. Since it all amounts to scripting the collection tasks, it also eases the development of a crowd-centric application that uses location providers and network sensors to detect densities and flows, or audio samples to estimate congestions. By default, the framework preserves the anonymity of the users and devices involved in the collection campaign, but the users can choose to reveal their identity, and in these cases we can collect social traits of the crowds such as gender and age distributions. Unlike most app-driven solutions we have analyzed, Medusa is a standalone, open-source, and ready-to-use framework with both client-side and cloud components.

The sensing application developed and employed by Wirz et al. [83] is integrated with a back-end framework for storing and processing collected data. This framework, *Coenosense* [82], receives location updates from the application, and was actually designed and employed for sensing crowds. It supports only the location-provider-sensing modality and functions in a straightforward way, without initiating sensing tasks or participant recruitment. The sensing application is responsible for enforcing the collection and privacy policies, while the framework receives anonymized samples for storage and real-time processing. The latter includes visualization, providing heat maps on crowd pressure (available only to the event managers). The processing and visualization mechanisms work only with aggregated location updates, restricting its usage to applications that collect these. Unlike Medusa, Coenosense is not open source and not freely available for download.

Rachuri et al. [61] propose *METIS*, an adaptive platform that offers support for offloading the sensing tasks of social-sensing applications. In METIS, sensing offloading is made possible by the existence of a sensing infrastructure in addition to the mobile application. The system is designed for detecting interactions between its users by combining audio recordings, Bluetooth-based proximity detections, and motion sensor data. Since some sensing modalities are more energy consuming than others, the system can distribute some of the sensing tasks to the sensors already placed in the environment. The overall goal of this platform is to reduce energy consumption and make the application as light as possible. To illustrate, the authors show that energy consumption can be very close to that with just using the phone without sensing and the WiFi on. Even though METIS provides a significant optimization of the energy consumption, the fact that it relies on networks of external sensors poses a major disadvantage.

The nature of the applications for which METIS is designed requires privacy controls and policies both on the client and at the back end, but the authors do not address the privacy issues. Their research goal goes beyond mere indoor detection of crowd patterns (groups, interactions between groups) by analyzing the data based on user profile and membership to certain teams and communities. In their field experiment, they use METIS as a surveillance platform to determine how often the users interact within their group and with colleagues from other projects.

Mori et al. [51] provide a more generic approach for sensing applications, inspired from their work with wireless sensor networks. They offer both client-side and server-side support for creating and managing sensing tasks. One of the key ideas of its design is the collaboration between the nodes, which makes it very suitable to crowd-centric sensing. Like Medusa, it offers a description language for creating sensing queries (tasks) but with a different distribution model. While its design is promising, especially its support for interdevice communication using radio modalities, large-scale testing and deployment have not yet taken place.

The reason we consider crowdsourcing frameworks in our discussion is their support for various sensing modalities and device discovery. Using the former, we can aggregate data such as locations and use them for analyzing crowd properties instead of individuals. The latter, when supported, enhances the role of the devices: obtaining data about the presence of other proximal devices.

A framework primarily focused on discovering and managing devices is Crowdwatch [40]. It combines high-power radio and low-power radio in a hierarchical architecture for discovering participant devices and selecting them for the data collection process. The framework has so far been evaluated just in simulation but never deployed. The authors do not properly evaluate crowd dynamics, but only briefly mention the discovery latencies. By and large, the system seems designed for wireless sensor networks rather than for a system using smartphones. It is debatable whether this approach truly offers advantages. For wireless sensor networks, it is relatively straightforward to estimate the energy savings of the devices when using this hierarchical communication scheme, especially when they run only this application. For mobile devices, these savings are much harder to assess, considering the fact that other applications may need Internet connectivity, so it is already enabled, or the user has the habit of having the WiFi always turned on. Moreover, the authors do not consider the side effects of their discovery protocols, the fact that switching off the WiFi or Bluetooth interfaces would affect the other applications using them.

Bakht et al. [5] propose CQuest, another solution that combines low-power and high-power radios for opportunistic discovery and cooperation between nodes, focused on energy efficiency. Unlike Crowdwatch, it was not only tested in simulation but also deployed on a small testbed of rooted Android phones, which revealed several challenges. In the implementation, they needed to adapt the scheme to the Bluetooth interface's restrictions, such as the lack of support for broadcast.

Diverging from the centralized model of the previous frameworks, Xiao et al. [86] claim that the current approaches for sensing applications that harness the power of crowds do not scale well with thousands or more participants. Under the assumption that the heterogeneities of mobile platforms place great stress on the development and deployment phases, the authors propose a system relying on virtualization. They use a proxy virtual machine for each device, which handles the data processing and the communication with the virtual machines of each application (one per user), all residing in the cloud. Such an approach has advantages in terms of usability and privacy, the users installing only one crowd-sensing service instead of multiple applications and having their data processed and stored in their own virtual machines. The authors do not discuss how well the system performs and deals with privacy when it comes to aggregating data from all its users. Like Crowdwatch, this system is not yet implemented.

3.5 Applications

Cenceme [49] is one of the first participatory applications specifically designed to support multiple sensing modalities. While the platform on which it was implemented is obsolete, its features and the entire design and evaluation approach are still relevant. They perform extensive tests not only on power and resource consumption but also on the impact of various factors on the sensing results. The integration of five sensing modalities and the modular design are the strong points of this system. Privacy and scalability are not very clearly addressed, although privacy is considered in its storage policies. Raw audio and acceleration samples are stored locally first until they are processed, after which results are uploaded to the server, together with the device's locations and Bluetooth addresses of discovered devices. It is not clear whether communication is secured or if scanned addresses are hashed. Scalability both on the client side and at the back end is not discussed.

A follow-up, VibN [50], was designed for sensing crowd densities and presenting in real time the available hot spots. This Live Points of Interest feature is the main incentive for users to share not only their location but also audio samples.

Crowdsense@place [14] is a more recent system, similar to VibN in terms of purpose and use of modalities. It also provides crowd density information on points of interest, but has a different data collection and processing approach. VibN aimed at collecting some user demographics and basic daily usage patterns. In contrast, Crowdsense@place collected much more data, aiming at identification of popular places, visit patterns, the way the application was used, and in which contexts the data was collected and shared.

Citysense [44] analyzes in real time information about points of interest, in particular nightlife attractions such as restaurants and clubs, and presents the users a map of busy places. Both the application (Citysense) and the platform used for collecting and aggregating location data (Sense Networks Macrosense) are no longer available, and accompanying research appears to have been discontinued. Density analyses and privacy policies were their distinguishing points. Its clear focus and tight integration of functionalities presumably contributed to its popularity and the little need for built-in incentives.

Early on, Kjærgaard et al. [37] proposed EnTracked, a system designed to manage several sensing modalities in an energy-efficient manner in order to track individuals. While the purpose of this system is not directly connected to crowd sensing, its sensor management mechanisms and application logic are still relevant. Entracked's design was closely coupled to the mobile platform used in experiments, a platform no longer available today. A new EnTracked$_{RT}$ version was also implemented in Android and used in [7]. The newer system has more sensor management strategies and performs better (more energy efficient and robust) than the previous one on both the Android platform and the older Nokia one.

4 INFRASTRUCTURE-BASED SENSING

In addition to using essentially on-body sensors such as mobile phones, systems for sensing a crowd can also consist of sensors placed external to crowd members. We refer to these systems as being infrastructure based. They rely mostly on statically placed sensors that vary from custom devices to WiFi routers or standard computers. As we have discussed in the previous section, for application-driven systems, it is challenging to come to a user base that can gather enough relevant and sufficiently accurate data. In contrast, for infrastructure-based systems, participants have little to no interaction with the sensors or knowledge of the sensing campaign. Nonetheless, they have their own challenges regarding the quality and relevance of acquired data.

In the discussions and classifications performed in this section, we consider several types of papers:

- **Spot-on**: the papers describing systems specifically designed for collecting data about crowds. They present both the collection mechanism and resulting datasets but also statistics and visualizations for describing a crowd's state.
- **Hybrid**: they are spot-on for crowdsensing but rely on both an infrastructure of nodes (usually WiFi access points) and a mobile application. Jamil et al. [31], Kjærgaard et al. [35], Kjærgaard et al. [38], Kjærgaard et al. [39], and Kjærgaard and Blunck [34] present such systems.
- **Related**: similar sensing mechanisms as the spot-on papers, having the potential of being used for crowd sensing, but having a slightly different target domain.

We also consider systems related in terms of architecture but that focused more on other aspects. For example, some papers consider proof of concepts for radio-based capabilities. Others focus on novel localization and tracking techniques. In one case, a Bluetooth-based system designed for urban traffic monitoring with sensors placed on traffic lights and lamp posts was also able to collect data about the crowds of pedestrians [58].

O'Neill et al. [57] and Nicolai and Kenn [55] are among the first using Bluetooth for crowd sensing, in particular measuring the fraction of detectable devices. This type of measurement is of interest also for systems dedicated to indoor commercial venues, such as that of Phua et al. [59]. The latter study the feasibility of Bluetooth for acquiring data on shopping behavior. They detected that over 30% of all devices had Bluetooth enabled and were able to determine the average visit duration and even correlate demographics to having Bluetooth enabled or not. Takafuji et al.

[72], Wada et al. [76], and Zhao and Shibasaki [87] use laser-range scanners for indoor tracking and localization. The first system using WiFi signal-strength measurements for indoor tracking was proposed by Bahl and Padmanabhan [4]. Their positioning accuracy was further improved by studies such as Evennou and Marx [21]. Rouveyrol et al. [64] demonstrate the ease by which WiFi routers can be infected to track individuals in a stealthy and light way.

Roggen et al. [63] and Wirz et al. [84] used on-body sensing devices equipped with accelerometers in order to study behavioral properties of a crowd. They used movement classifiers for detecting both individual activities and collective behavior: group formation detection and group detection. Although these systems are built for determining some characteristics of the crowd, their contributions lie more in the analysis part than in the sensing. The experiments were performed in small indoor areas using a few volunteers equipped with sensors. While the ground truth is easier to obtain in such scenarios, they are far from a wide, outdoor area deployment scenario. Moreover, it would be easier to appeal to a larger user base by using smartphones or wearables such as smartwatches instead of their custom sensing devices placed on a participant's leg.

4.1 Architecture

Infrastructure-based sensing systems, like most application-driven systems, have a centralized design. However, key design issues for application-based solutions are less relevant when an infrastructure is in place. For example, application-driven systems offer various types of policies for data collection, storage, processing, and communication, driven by energy efficiency, resource consumption, and privacy considerations. In infrastructures with static sensing devices connected to a power source, energy efficiency is no longer an issue. Likewise, ensuring privacy becomes generally easier for the simple reason that there is no application on the smartphone that needs to be trusted when it comes to crowd sensing. (Nevertheless, it is still surprising to see how much sensitive information is being leaked even by standard protocols [6].)

Most of the systems discussed in this section rely on static nodes that detect devices in their proximity. We also considered as infrastructure based the systems that employed mobile phones carried by volunteers. The solutions that fit into this category are those that do not concentrate on the application but on the collection process, and they provided very little information about the software running on the devices [54, 78, 79]. On the other hand, we consider solutions such as Chon et al. [13] to be application driven as they focused on the application's implementation, its functionalities, and its user interface and then tested it using volunteers.

Two of the systems also relied on badges, Bluetooth LE ones in Jamil et al. [31] and WiFi ones in Acer et al. [2]. For the latter, the choice of using badges was motivated by event-specific analysis purposes. The system used fixed WiFi scanners and collected two datasets, one with the mobile devices they detected and one with the badges provided to certain categories of participants.

The majority of the infrastructure-based systems in our survey use the sensing devices just for collecting data and uploading it to the server. We observe little variety in their policies. *Sensing* is enabled at deployment time and generally performed at fixed sampling rates, without the need for triggering tactics regarding sensing or communication: context enabled or demand driven (the server issues collection tasks or relays tasks provided by other users). Note that when relying on mobile phones for sensing behaviors, the system is dependent on the messages that are sent by the phones, which may be done at highly irregular intervals. *Processing* is performed at the server, mostly after a sensing campaign. *Communication* between the server and sensors is performed continuously, and generally data is stored at the server. In application-driven systems, the client-side *storage* aspect has a significant role mostly due to privacy issues. Depending on the application's features, the use of local storage can minimize the interactions with the server, for instance, when users visualize a track of their locations in the last 2 hours. In infrastructure-based systems, storage policies are usually dictated by the hardware design and software implementation choices of the sensing devices. Moreover, the device's main role is to merely sense the presence of the crowd and not to provide feedback to its owner.

4.2 Sensing Modalities

In Section 2, we presented the main sensing modalities used for collecting data about crowds. Mobile phones generally use the location provider, but also other modalities when energy consumption is at stake. In infrastructure-based systems, all spot-on solutions use only proximity detectors.

Unlike ranging sensors such as lasers or external motion detection sensors, the radio-based sensors do not actually detect a person's presence but rather his or her devices. Bluetooth was the most common technology employed before the growth of the smartphone market share. Currently, due to the limitations imposed on the Bluetooth interface by the phone manufacturers, the increase in WiFi usage and the widespread availability of hotspots in outdoor environments, we see a strong shift toward using WiFi signals for detecting devices. It is unclear whether this trend will persist, yet combining Bluetooth and WiFi systems seems a viable solution.

Many sensing infrastructures are designed for indoor environments. Indoor sensing systems are less related to crowds: they focus mainly on positioning and counting users. However, we identified some solutions [23, 24, 32, 65] designed for larger indoor venues and for detecting crowd movements and patterns. Regarding sensing modalities, indoor solutions prefer Bluetooth, laser ranging, or RFID, while only WiFi and Bluetooth are used outdoor. Indoor solutions can also be application driven or hybrid, detecting the hotspots placed in the building [35, 38, 61].

4.3 Crowd Properties

We observed that the surveyed systems approach the sensing layer both in a top-down and in a bottom-up fashion. With top-down approaches, which crowd properties need to be obtained are generally well defined, and appropriate choices for the employed technologies are made. In the bottom-up approaches, the sensing modalities and overall infrastructure are put to test. The system is evaluated based on the crowd properties it can sense. Regardless of the approaches, all surveyed systems describe the crowd's state through spatiotemporal characteristics. Some also infer behavioral primitives and social information. We look for dimensional properties (count, size, density), movement properties (flow, routes, speed), and social properties (behaviors, activities).

Some systems also look into individual tracks, a property not related to the crowd state, but which can be easily extracted from the dataset collected by the proximity detectors. Most of the systems (28 out of 34) detect counts, while less than half detect densities (13) or flow (11). Routes and speed are rarely addressed (just three and five, respectively). In Table 4, to be found in the appendix, we have marked with a *t* in the *routes* column the systems that are limited just to individual tracks and do not further analyze their aggregates.

While dynamics are the first properties that come to mind when describing or analyzing crowds, social aspects might be of interest as well, such as demographics, grouping, and other relationships. The social relations and demographic data can be deduced either in a privacy-intrusive way, as in Barbera et al. [6], or by employing noncomputational means such as questionnaires or purposefully selecting the participants in the sensing experiment as in Jamil et al. [31]. Out of the surveyed systems, 14 addressed social characteristics, most of them identifying stay durations and commuting patterns.

The system presented by Barbera et al. [6] sniffs probe requests and retrieves not only the MAC addresses but also the preferred network lists (PNLs), allowing one to derive social structures. It is possible to infer relationships between people based on the networks they shared and the types of those networks. The latter is deduced from the name of a network (e.g., revealing that it is a workplace, a public place, a cafe, etc.). The authors also analyzed the social influence of the vendor adoption by correlating the distribution of mobile-phone vendors with social relationships. This collection campaign raises privacy concerns since the authors placed laptops in certain locations and sniffed packets without removing the information that can identify and track individuals. The sensing mechanism employed by Jamil et al. [31] raises fewer privacy concerns. During a large outdoor festival, over 700 Bluetooth Low-Energy (BLE) tags and 24 smartphones were handed out to volunteers from various social groups. The authors achieved 80% accuracy in detecting groups, and performed group analysis by looking at routes, visit durations for certain attractions and cross-interactions, flows, and cohesion. They also mined community-related demographic data using the gender and age information collected from the volunteers.

4.4 Privacy

As discussed in Section 2, one of the functional features we look for in crowd-sensing systems is privacy policies. Most application-driven systems offer a form of privacy control through either user settings, secure communication, or anonymous data collection. The topic of privacy is also thoroughly covered in participatory sensing studies [15, 16, 67–69].

The proximity-based systems mode of operation raises more ethical concerns than those using other modalities (e.g., location-based applications) since participants are generally not aware of the collection campaign. One may consider that the tracking risk is an issue just in infrastructure-based crowd-sensing systems, but it is also in application-driven systems. Chon et al. [13] observed this by using just 25 users walking around Seoul for 7 weeks. They used an application that detected packets from surrounding phones and they were able to track a few devices for more than 8 hours a day.

The papers we surveyed describing infrastructure-based systems for sensing the crowds address privacy issues only briefly. The privacy requirements we consider for these systems are to not leak personal data and to prevent identification and tracking of individual users. These led to functional requirements such as secure transmissions, secure server-side storage, data collection restrictions (collect the minimum amount of data necessary for extracting characteristics about the crowds), and anonymization of device information.

The systems we surveyed use proximity-driven modalities for detecting mobile devices. The minimal information needed is a device identifier (typically its MAC address) and a detection timestamp. Less than half of the solutions we surveyed performed anonymization of the address. Others [75] considered that not using the device's name is enough for preserving privacy. Another concerning case is the WiFiPi experiment [11], in which the event organizers had access in real time to a web dashboard with the list of all detected addresses and the devices' manufacturers. Some systems [6, 57, 59] had no concerns about privacy, using the device names, addresses, or their SSID lists to extract social characteristics of the crowd (relationships, demographics, vendor distribution).

In Table 5, to be found in the appendix, we summarize the security measures applied in the collection, communication, and storage phases. We also include some architectural policies.

Some systems store data just locally to be extracted afterward. We observed that few systems used secure transmissions for data uploads to the server. Those that performed hashes, either on the sensing devices or on the server, did not discuss any other security issues or if their measures are enough to protect privacy.

Demir et al. [17] have analyzed the privacy policies of commercial WiFi tracking solutions from 15 major companies. They looked into their policies for data collection, data transfer, data anonymization, storage, data retention, and opt-out; only two of the solutions covered all these policies. Most of the companies employed some form of hashing for the MAC addresses they collected, but their methods were too weak and could be broken in a few minutes using just one high-end GPU. Just five companies used secure connections and had data retention policies. An option we have not seen in the academic projects we have surveyed is the opt-out. This permits the users to enter their MAC addresses if they are not willing to be tracked. Most of the commercial solutions had support for this option.

Recently, the OS vendors, including those for mobile devices, have included MAC address randomization support in order to prevent tracking. Windows 10, for instance, uses a random MAC address for each new network it connects to and reuses it when it connects again. Even though crowd-sensing systems do not intend to track individuals, they rely on the unique IDs of the devices in order to analyze the flow between certain locations. If the addresses are randomized, even just for the scanning phase, it affects the density and count analysis, and hinders the overall flow analysis, finding the popular routes and so on. A recent study by Vanhoef et al. [74] shows that tracking is possible even with these randomization mechanisms in place by leveraging the contents of the probe requests.

4.5 WiFi-Based Systems

Many proximity-based modalities employ WiFi technology. The IEEE 802.11 standard defines two scanning modes, to which we refer as passive and active scanning. In the former, the access points periodically announce their presence and the devices listen to them. These beacons are sent approximately every 100ms on only one channel, so devices must listen to different channels. Such a scanning technique is not very adequate for high-mobility scenarios like monitoring pedestrian traffic in urban areas. In active mode, it is the device that periodically scans for access points. This mode is more appropriate for crowd sensing, and all solutions we surveyed rely on it. IEEE 802.11 defines three types of frames: control, management, and data. Active scanning relies on management frames, called *probe requests*, transmitted by the client devices to discover access points (APs). Probe requests are sent regularly, at intervals up to 120 seconds, depending on factors such as vendor or power state. Bonné et al. [11], Maier et al. [47], and Musa and Eriksson [53] provide a more detailed overview of this detection mechanism. Surveyed papers refer to sensing devices as WiFi scanners, monitors, or detectors, and we will use these terms interchangeably.

This scanning process and its threats to user privacy is an active research topic. Recent experimental studies such as those conducted by Freudiger [22] quantify the privacy risk and assess the address randomization effectiveness. The probe requests contain the MAC address of the sender and, optionally, the SSID of the AP he or she wants to associate with. The sensing devices just receive these messages and retrieve the address. In addition to the MAC address, some systems also collect signal-strength measurements for better localization of the devices. While only one of our surveyed solutions also retrieves the SSID information, the fact that the probe requests contain this information raises serious privacy concerns, as explained by Lindqvist et al. [43].

Two of the spot-on systems conducted experiments utilizing both WiFi and Bluetooth, and also provided a comprehensive description of the scanning strategies. Schauer et al. [66] compared the accuracy and detection rate of both technologies in an airport, having access to ground-truth data from the boarding-pass scans. Their evaluation shows that WiFi is superior to Bluetooth for approximating the densities and the flow of the crowd, and has a considerably higher detection rate (4% Bluetooth/WiFi ratio). Abedi et al. [1] had a different approach, focusing more on the characteristics of these technologies. They tested the discovery time, detection range, and signal strength in various conditions, varying the environmental interference, the antennas they used, and the scanners' placement in overlapping regions. Their experiments were conducted at a smaller scale than Schauer's and the analysis was limited to counting detected devices and classifying people based on their movement speed (cyclist, runner, walker). Nonetheless, the observed detection rate between the two technologies was similar.

Several factors influence the quality of the collected dataset and its suitability for dimension and movement properties: antenna range, device transmission power, environment interference (objects, walls, number of people), probe request transmission rate and subsequently the discovery time, and scanner and device placement. Some of them can be addressed in the pre-experiment phases through either simulations or small-scale empirical tests. An example of the latter is the probe-request transmission rate, which varies between 1s and 120s. This rate depends on the monitored area [27], vendors, power states, screen status, or if the device is already connected to an AP.

Discovering the transmission rate is not as easy as it would seem. One experiment conducted by Li et al. [42] showed that iOS devices have a 70s to 1,200s interval, depending on the device's state, Windows devices varied between 10s and 1,200s, and devices running Android between 1s and 2s. In contrast, Fukuzaki et al. [23] conducted similar tests and found an approximate period of 480s for iOS and values between 15s and 250s for Android. On the other hand, Schauer et al. [66] had observed in their tests that iOS devices send probe requests more frequently than Android devices. These inconsistencies may be attributed to differences in the API versions of the tested devices. We stress the fact that we have a heterogeneity not only in the devices' underlying hardware but also in the API versions they use, especially for Android devices. Pre-experiment tests should be performed for evaluating the sensing characteristics (discoverability, detection range) using various device models and software versions.

One of the criteria in any infrastructure-based sensing system is the efficient placement of scanners and sensors for covering an area. For the systems designed for sensing movement properties such as flow or routes, this task is particularly challenging. In an urban environment, having the sensors placed too sparsely considerably diminishes the accuracy for determining routes, since pedestrians may take several paths to get from one point of interest to another. Even though it has lower deployment and maintenance costs, spatial sparsity is not a viable option when systems also collect signal-strength measurements for performing localization using triangularization. On the other hand, having a denser network of sensors generates overlaps, which means that the device is detected by two or more scanners at the same time. Abedi et al. [1] see this as an opportunity to assess the movement type based on the time passed in the overlapping region, while Musa and Eriksson [53] consider it as a negative factor influencing the system's tracking accuracy.

The dataset can be affected by temporal factors; for instance, the person may pass the monitored area without the device transmitting any probe requests, or the device's state and placement at that time reduce its transmission power. Also, the WiFi scanners and monitors may suffer unpredictable downtime, an issue rarely discussed in our surveyed papers and which notably affects the short-term and medium-term deployments. Bonné et al. [11] mentioned the technical difficulties they encountered during their experiments, which included even power failures.

Musa and Eriksson [53] offer one of the most relevant papers on the challenges regarding tracking and also provide the only infrastructure-based system that considers a form of incentives. Their three experiments were well tailored to the type of analysis they wanted to conduct. In order to enhance the chance of detections, they considered three methods, two of them using AP emulation and one using RTS injections. The first two increase the number of detected devices and the latter the number of packets received from each of them. We see these as incentive mechanisms transparent to the monitored participants, since they encourage the devices to associate with their monitors. They advertise their monitors either as popular SSIDs or as SSIDs with which the devices have connected in the past. All three techniques were tested, with AP emulation giving the best results.

The hybrid systems of Kjærgaard et al. [35, 38, 39] and Kjærgaard and Blunck [34] rely on a WiFi infrastructure for positioning. The mobile applications detect the APs in their proximity and record the signal strength and send the data to a centralized server. A WiFi fingerprinting step must be performed prior to the experiments, which highly affects the scalability of these systems. In the case of Kjærgaard et al. [35, 38], the applications also collect motion sensor data for a multimodal analysis. They apply classification algorithms for detecting flocking behavior [38, 39] and leadership and following patterns [34, 35]. The systems performed very well, having good detection accuracy, but the tests were conducted at a smaller scale, in indoor environments with only a few volunteers (10 to 19).

In Table 6, to be found in the appendix, we have summarized the experiments of the WiFi-based solutions we surveyed, which, with the exception of Rouveyrol et al. [64] and Bahl and Padmanabhan [4], are all spot-on when it comes to sensing crowds. We classified the duration of the experiments and the size of the monitored areas based on the observations for all the systems we surveyed, including the participatory ones. Although for the app-driven experiments we also had applications publicly deployed for many months, for the rest of them, the time intervals and the area types had similar variations. None of the surveyed sensing infrastructures was a permanent one, even though we suspect that some [52, 65] remained in place even after the period covered in the analyzed dataset. The duration of the experiment can be *short term* (ST) if it lasts less than 24 hours, *medium term* (MT) when lasting between 24 hours and 7 days, and *long term* when lasting for more than 7 days. Few deployments lasted for more than 1 month. The environment is classified as being *small indoor* (SI) for less than 100m^2, *medium indoor* (MI) when dealing with a floor or a small shopping area, and *large indoor* (LI) for several floors or even buildings, exhibition halls, and shopping malls. For outdoor events, we distinguish between *small outdoor* (SO), such as a playground; *medium outdoor* (MO), for example, when dealing with a relatively small campus; and *large outdoor* (LO) for mass events and entire city centers.

Although no information is given in the papers in terms of deployment and maintenance costs, we have observed that they employed low-cost customized solutions based on either Raspberry Pi or low-end access points [53].

4.6 Bluetooth-Based Systems

Bluetooth is a well-established standard for proximity-sensing systems. The broad use of Bluetooth in consumer electronics, notably personal devices, makes it ideal for applications that sense or track human presence. Since in our survey we focus on sensing crowds, we looked more into Bluetooth-based systems for large, crowded, outdoor areas than for indoor venues or building surveillance.

In order to communicate, Bluetooth devices pass through several steps, but for sensing purposes, we need just the inquiry phase. In this step, a device referred to as *master* sends inquiry requests and all nearby listening devices respond. The response contains the MAC address and some additional information such as the name and class of the device. The class can be used in filtering the nonmobile devices. Master devices deployed in the sensing infrastructure just detect passing devices through the inquiry phase and do not establish connections with them.

The problem with this sensing mechanism is that the phones reply to inquiry requests only when they are in *discoverable* mode; it is not sufficient to just have the Bluetooth interface enabled. Due to security concerns, the mobile OS vendors have imposed certain restrictions, and current smartphones have a limited discoverability window (less than 5 minutes) and users must explicitly consent to it. The latest Bluetooth experiments in crowd sensing, performed in 2014, show a very low percentage of discoverable devices (under 10%) as opposed to WiFi devices. What is interesting, though, is that we do not see a significant drop between older studies and newer ones for this percentage. To what extent wearables are going to change this situation remains to be seen (wearables generally require a Bluetooth connection to a smartphone).

Bluetooth devices can communicate at a range between 10m and 100m, usually 10m for mobile devices, more for laptops. With a lower range than WiFi, Bluetooth-based infrastructures must rely on a higher density of sensing devices, which can increase the deployment and maintenance costs. As in the case of WiFi, the sensors must be connected to a power source, and preferably to an Internet connection for uploading the collected data. When covering large areas without wired Internet connectivity, WiFi-based systems rely on mobile data connections, as in the case of Bonné et al. [11], in which a phone was attached to every sensing node, 15 in total. Such an approach may not scale well in terms of price for Bluetooth infrastructures. Some Bluetooth-based systems log data locally on sensors and do not upload it in real time to a server (see Table 5 in the appendix).

The first Bluetooth-based crowd-sensing systems focused on its suitability for detecting people, notably the rate of discovered devices. Nicolai and Kenn [55] performed a field study in crowded commercial areas and measured rates between 2% and 6%. Unfortunately, they do not discuss the results in respect to the phone-market penetration at that time, in 2006.

The trend of transitioning from Bluetooth to WiFi for crowd sensing can be observed in the conclusions of several Bluetooth-based papers, not only in the comparative studies of Abedi et al. [1] and Schauer et al. [66]. Phua et al. [59] conducted a questionnaire asking the participants whether or not they have both Bluetooth and WiFi enabled and whether they would use a free WiFi public hotspot provided in the venue. The responses were in favor of WiFi.

We can also see a transition to Bluetooth in the experiments conducted by some researchers, for example, in the case of Weppener et al. They relied on Bluetooth-based systems [78–80] and in [77] they deployed a WiFi-sensing infrastructure during a 2015 mass event. In their latest approach, they refer to WiFi and Bluetooth together in the first part of the paper, while the sensing campaign focuses solely on WiFi. Even though the data is collected in a similar fashion, they seem to ignore the differences between these interfaces when it comes to being able to discover devices. It seems improbable that using only Bluetooth, they would have been able to detect two-thirds of the event participants, like they did in their WiFi experiment.

BLE is another technology for proximity sensing, but due to its novelty, it is still less prevalent in crowd-sensing systems. Jamil et al. [31] discuss the only surveyed system that uses it. The authors obtained high accuracy in discovering tags given to hundreds of participants. BLE has a range similar to classic Bluetooth but a different mechanism for discovery and communication, leading to lower latencies and better discoverability.

Since many of the surveyed Bluetooth papers focused on the detection rate, they also provide clear statistics about it. In contrast, only three of the WiFi papers offered this information: Bonné et al. [11] (29.3%), Fukuzaki et al. [23] (30%), and Musa and Eriksson [53] (68%). Details can be found in Table 7 in the appendix.

4.7 A Few Notes on Ground Truth and Visualizations

The infrastructure-based papers we surveyed present instrumented or noninstrumented experiments. In the first case, we have small-scale lab experiments, usually indoor, with volunteers following certain scenarios. They have the benefit of providing ground-truth data in an easy manner, sometimes just by manual observations. Such experiments are also suitable for the predeployment phases we discussed earlier or for testing certain analysis methods. Unfortunately, noninstrumented experiments, especially those conducted at a larger scale, are hard to validate in the absence of another source of data (e.g., CCTV video streams). In some cases, the researchers can leverage venue-specific information, such as airport security data [20, 66] or turnstiles (as done by Ellersiek et al. [20]).

For large areas, some systems relied on short-term data acquisition for validation purposes. For instance, Versichele et al. [75] computed the detection ratio during 10 video-based experiments, each lasting up to 15 minutes. The systems relying on volunteers for observing the participants also collected validation data during a limited period of time. Kalogianni et al. [32] designed a people-counting application and also surveyed random participants. Musa and Eriksson [53] traveled in the monitored area, recorded the GPS traces of their devices, and determined the tracking accuracy for those devices. In other cases, such as Mueller et al. [52], validation was performed using data from other similar deployments.

The process of acquiring ground truth can be particularly labor intensive, as in the case of Weppner et al. [77]. Over 40,000 annotations needed to be *manually* processed from 71 ground-truth images collected from a camera covering almost the whole monitored area during the entire duration of the event.

It may be questioned to what extent some ground-truth collection methods are actually valid. Naini et al. [54] validate their population-size estimations using data collected at the entrances and exits of the event premises. Since their choice of technology was Bluetooth, they placed phones that count the number of Bluetooth discoverable devices. The problem is that a phone visible at the entrance may not be visible 2 minutes later or the other way around.

Real-time visualizations and analysis were less addressed in the infrastructure-based systems than in the application-driven ones. Only three spot-on systems included it [11, 24, 42], and a few others had a design that could support it. None of these systems mentioned real-time feedback to the public through websites or dedicated applications. In other sensing domains, such as pollution tracking, sharing the results with the community and attracting the interest of the public is one of the priorities [18].

5 CROWD-CENTRIC SENSING SYSTEMS

Thirty of the application-driven and infrastructure-based sensing systems we have discussed in the previous sections are very relevant for sensing the crowds due to the quality of their collection campaigns and the relevance of the dataset. With few exceptions [2, 11], they also include the analysis of the crowd dynamics. We consider them spot-on to the purpose of our survey, as previously argued in Section 2.

We organized these solutions based on their purpose and employment and their underlying architecture.

- Monitoring crowds during *mass events* usually taking place in a short period of time, having tens of thousands of participants and covering a large area. The majority of systems were designed for this purpose (six application-driven and 12 infrastructure-driven solutions).
- Monitoring crowds of pedestrians in *cities* (usually city centers or pedestrian areas) during larger periods of time (four application-driven and seven infrastructure-based solutions). In two of the cases [6, 12], mass events took place during the monitoring period.
- Infrastructure systems deployed for monitoring crowds in large *indoor* areas (seven solutions, four of them during mass events).

Table 8 lists them based on their purpose and design. An ideal system for sensing crowds provides data and communication privacy to the monitored participants and incentives to attract a large user base. It also scales in both area size and density of participants without a considerable increase in development, deployment, and maintenance costs. The analysis methods support various density scenarios. The sensing application does not have a considerable impact on the device's

Table 2. Criteria for Comparing the Spot-On Papers

	Rating 0 (worst)	Rating 1	Rating 2	Rating 3 (best)
Privacy	No security and privacy mechanisms and policies; no discussion; e.g., [6, 79]	Some privacy but not enough + no discussion; no mechanism implemented + discussion; e.g., [11]	Standard measures; not a primary concern; e.g., [71]	Ensures privacy: anonymity, communication security; high concern; e.g., [23, 50]
Incentives	No incentives; no discussion; e.g., [56, 70]	Provides incentives; no details about them or effectiveness; e.g., [80]	Provides incentives; not very effective; not a primary concern; e.g., [71, 83]	Provides incentives; effective; research & tests on incentivization mechanisms; e.g., [9]
Scalability	Hard to extend the system to more users or cover a larger area; e.g., [38]	Scalable in some degree, more close to 0; e.g., [13]	Scalable, not a main focus of the paper or not directly discussed; e.g., [20]	Highly scalable; designed for scalability; discussed in the paper; e.g., [9, 50]
Transparency	The participants need to often interact with the system; e.g., [71]	Active participation is required, e.g., [50]: the participants start the media recordings	Unclear, but may be transparent, e.g., [56]: seems transparent but not discussed, it's based on sound recording so it might require permissions to trigger it	The participants are not required to do anything or they are not aware of the collection campaign; e.g., [48]
Resource consumption	High energy consumption, many resources involved; no sensor sampling strategies; e.g., [56]	Unclear, but most likely high, e.g., [48]: collected battery levels, observed that the battery dropped 20% in 2 h, controlled the submit rate based on the critical places	Unclear, but most likely low, e.g., [80]: uses Bluetooth and GPS but implemented dynamic sampling policies and geofencing	Efficient use of resources; e.g., [75]
Ease of deployment	Requires a lot of effort (technical, logistic, marketing) and high costs in equipment, development, maintenance; e.g., [14, 53]	Most likely demanding, e.g., [13]: they added kernel drivers, required rooted phones	Most likely not very demanding, e.g., [83]: developed an app based on an existing framework and collaborated with the event organizers for distributing it as the official event app	A simple app and a server rented in the cloud or using existing infrastructure and just obtaining the data; e.g., [65]

battery or affect other running apps. The users are aware of the data collection performed by the application and are in control of its sensing settings and permissions. Moreover, the usability of the application is not hindered by frequent requests for user input. All these traits are present in various degrees in the spot-on systems. Real systems need to trade some characteristics for others based on the crowd properties they want to assess, the level of accuracy, the sensing technologies, and the desired interaction with the users.

We use a rating system for six features relevant to crowd-sensing systems: privacy, incentives, scalability, ease of deployment, resource consumption, and transparency. The diagrams in this section depict the tradeoffs between the six features. Other characteristics such as the quality of the dataset, the coverage, the accuracy of the detections, or the usability of the applications either are not discussed in the papers or vary too much to have a common comparison ground for numeric ratings.

The ratings vary between 0 and 3 and are scaled to the worst and the best implementations we have seen in the surveyed systems. When distinguishing between the ratings, we also consider the concern of the authors for that particular topic. Table 2 summarizes the meaning of these ratings.

For transparency, resource consumption, and ease of deployment, the middle ratings depend on each system's implementation.

For infrastructure-based systems, resource consumption refers only to the scanned devices, and with only one exception [53], all the surveyed systems have the higher rating. Musa and Eriksson [53] employ mechanisms for increasing the packets sent by the devices, thus enhancing the number of detections.

The applications employed for events and urban sensing have more varied ratings for resource consumption. The system with the highest rating is proposed by Kannan et al. [33]. It is designed for counting and density, and achieves good accuracy with a very low power consumption. This system is the only one of the spot-on systems that employs a tone-counting mechanism. Unfortunately, the solution does not seem to scale to thousands of users.

Transparency refers to the interaction between the monitored participants and the system. We do not include data transparency since most of the papers do not discuss it or share any results or information on the collection campaign (only 14 out of over 90 systems we surveyed actually share any data with the participants).

We grouped the systems based on their purpose, modalities, and the monitored environment (indoor/outdoor) as reflected in Table 8. The application-driven ones had more diverse ratings, while for infrastructure and hybrid ones some of the criteria were constant. For indoors, all the spot-on systems relied on an infrastructure-based or hybrid architecture, and with no incentives and less interest in privacy, as shown in Figure 5 in the appendix.

For mass events, which consist of thousands of participants and cover large areas or have high densities, the applications were fairly easy to deploy, being less cumbersome to market as official apps than to make them available through official app stores. The latter often form a hindrance due to their various regulations (Weppner et al. [80] discuss some of these restrictions). On the other hand, the infrastructure-based systems were considerably more difficult to deploy as we show in Figures 3 and 4 in the appendix. They have lower ratings especially for events rather than for urban monitoring, mostly due to the monitored areas' limitations. Also, the lack of Internet connectivity forced the use of local storage, making deployment and data collection less scalable. The easiest deployment for infrastructure-based systems was in the case of the building monitoring performed by Ruiz-Ruiz et al. [65], in which they leveraged the existing and already dense Wi-Fi infrastructure.

The applications provided a more consistent enforcement of privacy than their infrastructure-based counterparts. Out of the 13 event-related infrastructure systems (both indoor and outdoor), only five had the highest rating, while the rest had 0 or 1, depending on their discussions on this subject.

Some of the spot-on solutions also collected ground-truth data, which permitted them to validate the datasets using relative error [80], correlation coefficient [83], tracking accuracy [42], or detection rate [31].

The spot-on systems with notable good accuracy, in-depth experiments (varied scenarios, multiple types of mobile devices), and complex analysis (e.g., considering energy consumption, scalability) are Kannan et al. [33] for crowd counting, Weppner and Lukowicz [79] for density estimation, Nishimura et al. [56] for congestion classification, and Kjærgaard et al. [38] for flock detection. Regarding the latter, the flocks and other patterns such as leadership and following are the topic of the systems described by Kjaergaard et al. in [34, 35, 38, 39]. Due to their similarities in the sensing system and the emphasis on the processing part, we included only one of these papers in the spot-on papers category. They provide high detection accuracy, especially when combining modalities—WiFi AP detections with motion sensors [39]. Also on flocks analysis, Wirz et al. [85] provide algorithms based on GPS traces to detect them with a high accuracy of around 78%. The dataset was collected using an instrumented experiment with volunteers carrying smartphones with dedicated sensor logging software, thus not making the focus of our spot-on papers classification.

In the absence of ground truth, the validation of the dataset is tricky and its quality is deduced based on the crowd properties it can describe, for instance, counts and densities. A notable mention is Blanke et al. [9], who capture the density and flows during a large festival with hundred of thousands of participants. Using the same dataset, the authors also obtain high map reconstruction accuracy in Blanke et al. [8], an analysis-oriented paper. Their experiment owes its success to the use of incentives and providing the official application of the event.

In some cases, the dataset is not enough and the researchers also rely on simulations. For instance, the application-driven system deployed by Stopczynski et al. [71] did not manage to attract enough users and they extrapolated the coverage through simulations. The authors discuss vital points regarding the challenges of their chosen technology (Bluetooth), compare the application-driven and infrastructure-based approaches, and are optimistic about Bluetooth's potential for crowd mobility sensing despite their aforementioned disadvantages. In the case of Mallah et al. [48], the spatiotemporal data was used just for checking the battery consumption, and the analysis part (group detection) relies on simulations for comparing clustering algorithms. This is also one of the papers that offers incomplete details on the sensing part, such as how they computed the positioning accuracy and what exactly were the incentives.

6 DISCUSSION AND CONCLUSIONS

Counting pedestrians has almost become commonplace, and using techniques such as WiFi scanning or Bluetooth tracking is gaining widespread popularity. However, we have been able to find only relatively few groups that report on real-world experiments with automatically sensing the behavior of crowds. This is somewhat surprising considering the open-ended issues that our study reveals. For one thing, we have not been able to identify any solution that adequately addresses all aspects of transparency, incentives, privacy, scalability, ease of deployment, and resource consumption. For example, while app-driven solutions are generally best at preserving privacy, they do require that users actually install a solution on their phone, in turn hindering practical scaling and requiring incentive mechanisms. Likewise, infrastructure-based systems have the advantage of transparency and scalability, but one may consider it troublesome to see how little attention is paid to handling privacy.

In practice, we see that large-scale solutions are being applied for gathering data from crowds. For example, many modern festivals use electronic bracelets that combine the function of a ticket, wallet, and tracking device. In many cases, tracking is limited due to the use of passive RFID technology, which requires readers that operate at relatively close distance. However, applying alternative technologies such as UHF RFID and active RFID allows reading at larger distances. Also, connecting RFID wristbands to smartphones opens up paths for tracking users, notably in combination with social media scanning. Unfortunately, up to this day there are no detailed reports on how these or other techniques are actually being deployed. Based on this survey, we can only suspect that even in these cases there is still much to improve, notably concerning privacy, but also technical issues such as accuracy, transparency, and resource consumption.

When it comes to accuracy, we have seen very mixed reports to the extent that we were not able to include it as a criterion for evaluation in Section 5. A major problem is acquiring ground truth. We have seen only few studies paying explicit attention to decently validating their results, which is in line with conclusions drawn by Wijermans et al. [81]. Despite the difficulties, we believe it is necessary to pay attention to this issue, if only using additional, independent measurements to identify anomalies in the original dataset.

It seems that only by combining app-driven and infrastructure-based solutions will we be able to come to decent solutions for sensing crowds. In essence, such solutions would combine the power of local measurements (i.e., by end-users) and global ones (through traditional scanning techniques). Privacy would be much in the hands of the user, although unintrusive scanning techniques do require that data is anonymized or even immediately processed to a sufficient aggregation level so that the original measurement can be subsequently discarded. Transparency in a hybrid solution can be addressed through techniques that are bundled upfront into the smartphone without requiring a special application to be downloaded. Instead, the user would simply need to configure his or her phone by switching certain services on or off. Furthermore, with further expansion of WiFi and Bluetooth in public and private places, the ease of deploying infrastructure-based solutions is certainly expected to improve, making a hybrid solution scenario even more plausible.

By and large, we conclude that automatically sensing crowds has already come a long way, but that there are significant challenges to be addressed before we can speak of satisfactory solutions.

APPENDIX

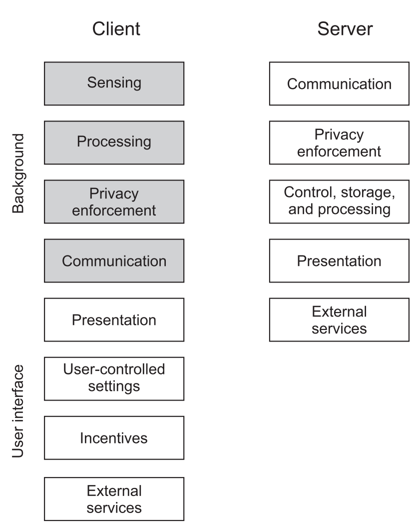


Fig. 2. General architecture for application-driven crowd-centric systems.

Table 3. Notable Frameworks, Applications, and Middleware Systems for Sensing Using the Crowds

System	Main Purpose	C	BE	Energy	Privacy	Status
Medusa	Sensing-driven crowdsourcing; task management and participant recruitment	Yes	Yes	Resource usage policies for low battery	Privacy controls worker anonymity	Available
Metis	Social context-aware sensing; sensing offloading			Offloads to a sensing infrastructure	Not considered	Implemented, not available
Coenosense	Crowd monitoring	No	Yes	No energy-aware strategies; high consumption due to GPS sampling rate	Anonymous data transfer, full user control over data	Implemented, not available
[51]	Sensing applications middleware; task description language			Efficient node selection	Not considered	Implemented, not available
Crowdwatch	Crowdsourcing framework; participant discovery	Yes	Yes	Considered in the evaluation, inconclusive results	Not considered	Not implemented
[86]	Sensing-driven crowdsourcing			Less communication on the device	Storage and processing users' containers	Not implemented
Cenceme	People-centric sensing; presence sharing	Yes	Yes	Power consumption benchmarks	Privacy controls	No longer available
VibN	Place-centric sensing; urban POIs			30min duty cycle	Secure communication, privacy controls, anonymized data	No longer available
Entracked	People-centric sensing; energy-efficient tracking	Yes	No	Dynamic sampling rate strategies	Not considered	Implemented, not available
Crowdsens @place	Place-centric sensing; visits per location			Dynamically adjusted sensor sampling	Privacy controls	Implemented, not available
Citysense	Place-centric sensing; urban POIs	Yes	Yes	Not considered	No details; data anonymity claims	No longer available

Not all of them are designed primarily for sensing crowd properties, but they could be employed for it too. **C** stands for *client*, **BE** for *back end.*

Table 4. Crowd State Properties Analyzed by the Papers Discussed in this Section

Solution	Dimension		Movement			Social	
	C	D	F	R	S	Group	Activity
[70]	✓	-	✓	-	-	-	Stay duration
[78]	✓	✓	-	-	-	-	-
[34, 35, 38, 39]	-	-	-	-	-	Flock detection	Leadership & following patterns
[53]	✓	-	-	t	✓	-	-
[75]	✓	✓	✓	-	-	-	Stay duration; returning visitors
[1]	✓	-	-	-	✓	-	-
[6]	✓	✓	-	-	-	-	Social relations; social structure
[11]	✓	✓	-	-	-	-	-
[20]	✓	-	✓	✓	-	-	Stay duration; walk duration
[79]	✓	✓	-	-	-	-	-
[23]	✓		✓	✓	-	-	Stay duration
[65]	✓	✓	✓[1]	-	-	-	User roles; stay duration
[66]	✓	✓	✓	-	-	-	-
[24]	✓	-	✓	-	-	-	-
[31]	✓	-	-	-	-	Detection; interactions; cohesion	Congestion; social structure; stay duration
[32]	✓	✓	✓[2]	-	-	-	User role[3]
[42]	✓	✓	✓[4]	t	✓	-	-
[52]	✓	✓	-	-	-	-	Commuter detection
[54]	✓	✓	-	-	-	-	-
[12]	✓	-	-	t	-	-	-
[77]	✓	✓	-	-	-	-	-
[57]	✓	-	-	-	-	-	-
[55]	✓	-	-	-	-	-	-
[?]	✓	-	✓	✓	-	-	Stay duration
[?]*	✓	-	✓	-	-	-	-
[2]	✓	✓	-	-	-	✓[5]	-
[59]	✓	-	-	t	-	-	Stay duration; demographics
[64]	-	-	-	t	-	-	-
[58]	✓	-	-	-	-	-	-
[84]	-	-	-	-	-	Detection	-
[63]	-	-	-	-	✓	Detection	Queuing
[72]	-	-	-	t	✓	-	-
[76]	-	-	-	t	-	-	-
[87]	✓	-	-	t	-	-	-
[4]	-	-	-	t	-	-	-

Spot-on and hybrid papers are first, ordered chronologically, followed by related papers grouped by their addressed topic. **C** stands for *count*, **D** for *density*, **F** for *flow*, **R** for *routes*, and **S** for *speed.*

Table 5. Summary of the Measures Employed for Preserving Privacy in Infrastructure-Based Systems. For the Systems Marked "–" We have no Information about the Respective Aspects

Solution	Sensing device	Communication	Server and Storage
[70]	Hashed MAC - SHA-256	Not supported	-
[?]	Hashed MAC - SHA-256	Not supported	-
[53]	No anonymization; transmit data every 1s	Support for no connectivity (data mule)	No anonymization
[75]	No anonymization	Not supported	-
[1]	No anonymization	Not supported	-
[6]	No anonymization	Not supported	-
[11]	No anonymization	No encryption; periodic uploads	No anonymization
[20]	Hashed MAC - SHA-256	Not supported	-
[?]*	Hashed MAC - SHA-256	Not supported	-
[79]	No anonymization	Not supported	-
[23]	Hashed MAC - SHA-1	SSL	Cloud
[65]	No anonymization	No encryption; triggered when detecting a device	Hashed MAC
[66]	No anonymization	Not supported	-
[24]	Hashed MAC - SHA-256	SSL	Cloud
[31]	Scans for BLE tags every 5min	Only when WiFi available	No need for anonymization
[32]	Hashed MAC	No encryption; periodic uploads	-
[42]	No anonymization	Only via WiFi; triggered when detecting a device	-
[54]	No anonymization	Not supported	Log files provided by APs admins
[52]	Hashed MAC - SHA-256	SSH tunnel PSK; periodic uploads	Institution-owned server
[2]	No anonymization	Not supported	-
[12]	No anonymization	-	Hashed MAC
[77]	-	-	Anonymized MAC, no info on the mechanism

Table 6. Details about the Experiments of WiFi-Based Infrastructure Sensing Systems

Solution	Timeframe & Environment	Sensors	Experiment & Dataset Details	Ground Truth	End Purpose
[53]	ST, SO	Off-the- shelf APs	(1) 5 nodes, campus, 9 months, 400K unique MACs. (2) 6 nodes, moderately busy roads, 12h, 20L unique MACs. (3) 7 nodes, 2.8km high-traffic road, 12h, 23K unique MACs; 68% mean probability of detection	GPS traces	Proof of concept: crowd tracking
[1]	ST, SI, & SO	Off-the-shelf WiFi, BT scanners	Popularity assessment experiments in 6 locations. 90% of observed unique IDs were WiFi addresses	No	Technology comparison for crowd monitoring
[6]	ST & LT, LO	Laptops, 1 fixed antenna	11M probes, 160K unique devices, 8 events, 6 experiments, lasting from 40 min to 6 weeks	No	Social study of the crowds
[11]	ST & LT, LO	RaspberryPI	(1) Festival, 3 days, 400m x 500m festival area, 138K unique devices. After filtering: 29K devices, 300K datapoints. 29.3% detected people. (2) Campus, 4 detectors, 3 months. 1,383 daily unique device; no analysis details	No	Tracking the visitors of mass events
[23]	MT, MI, & O	RaspberryPI	(1) 3 days indoor. (2) 1 day outdoor. No RSSI collection in either case.	Partially	Disaster prevention
[65]	LT, LI	Existing APs	15 days, 798 APs, 22 buildings, 10ha area. 10^9 measurements, 18K unique devices	No	Crowd behavior indoors
[66]	LT, MI	Laptops	16 days, 2 laptops, WiFi and BT scans. 6,211 daily unique WiFi devices, 250 for BT	Venue specific: airport gates data	Technology comparison for crowd sensing
[24]	LT, MI	Custom sensors	2 months, 20 sensors, shopping mall. 30% recognition rate	Motion detectors	Planning for commercial facilities
[32]	ST, LI	No info	20 monitors, 7 indoor campus areas, 1 week, detections every 10s. Outliers: outdoor or forgotten devices. Data from the monitors placed in 5 faculties, but was not good enough for analysis (low coverage). Ground truth not covering the whole period; no information on the questionnaire	People counting app; online questionnaire	Proof of concept: campus's rhythm
[42]	ST, SO	RaspberryPI	2 separate experiments. Walking type experiment: 14 nodes, 7 locations, 3 phones (1 per vendor). Density monitoring: 2 nodes, 30h	Partially	Proof of concept: crowd tracking
[2]	MT, LI	RaspberryPI, custom WiFi badges	3-day event, 40K visitors, 6,000m^2 venue, 30 gateways; 85 visitors with badges; badges' probe sending rate dynamically adjusted based on detected motion; after filtering: 2,526 unique devices out of 290K, 61 badges out of 85	No	Event crowd dynamics, behavior of two groups of people relevant to the event

(Continued)

Table 6. Continued

Solution	Timeframe & Environment	Sensors	Experiment & Dataset Details	Ground Truth	End Purpose
[12]	MT, LO	Custom sensors	27 sensors. Data from multiday festival in city center. Focus on mobile device detection, path detection	No	Effective data-cleaning techniques before analysis
[77]	LT, LI	No info	13-day mass event in 2015. 31 WiFi scanners, 9,000m^2. 209M probes, 85M after filtering, 300K unique devices. Mapped 2/3 of the visitors, 20% crowd density error. Additional localization experiment with 1 volunteer, 2 phones, 4.5m–5.6m mean localization error	Yes - video, 1 camera, manual annotations	Proof of concept: crowd monitoring using WiFi scanners
[4]	ST, LI	Off-the- shelf APs	Building floor with 50 rooms. 3 WiFi base stations. Validation using preliminary measurements and propagation model	Yes	RF-based indoor localization & tracking
[64]	SIM, LO	Off-the- shelf APs	Simulation based on real dataset. 1 volunteer, 1 device, 50h. Lab experiment with modified off-the-shelf routers. Performance overhead measurements	No	Proof of concept: tracking using modified routers

Table 7. Details about the Experiments of Bluetooth-Based Infrastructure Sensing Systems

Solution	Timeframe & Env.	Sensors	Experiment & Dataset Details	Detected Devices	Ground Truth	End Purpose
[57]	ST, SO	Laptop	1 laptop, 3 BT dongles; 10 gatecounts in the city, 30min each; 2 fixed gatecounts, long term; data from devices completely scanned	8%	Manual observations	Measure the percentage of discoverable devices
[55]	ST, MI & MO	1 laptop, 1 phone	4 locations—different types, cities, countries; 1–2h on different days	2%–6%	Manual observations	Measure fraction of discoverable devices
[70]	MT, LO	Custom	Festival, 48h, 40,000km^2 area; 1 measurement every 3s; 870K records, 12,700 unique devices; removed 55% of records due to cars	0.2%	No	Global crowd movements
[78]	MT, SO	Android phones (1 type)	3 volunteers each with 3 phones (2 in front pockets, 1 in back pocket); 3 days, 500m-long pedestrian zone, Oktoberfest2010; crowd density classification accuracy over 80%	-	Photos	Crowd density estimation
[75]	LT, LO	Custom sensors, 2 types	10 days, 22 locations; extracted 152K trajectories, 80K users; visual counts 10 times, for 15min each, in 8 locations	11 ± 1.8%	Visual counts	Bluetooth tracking for mass events

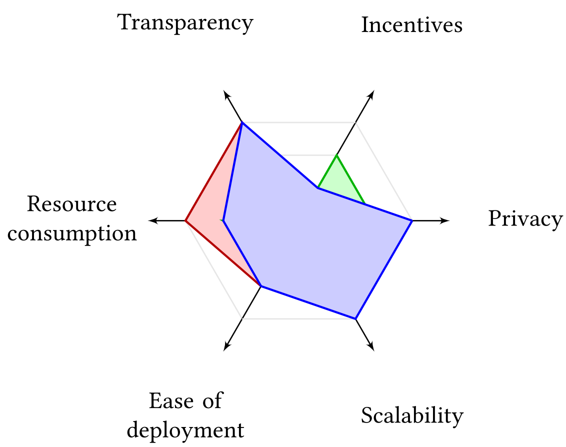
(Continued)

Table 7. Continued

Solution	Timeframe & Env.	Sensors	Experiment & Dataset Details	Detected Devices	Ground Truth	End Purpose
[1]	ST, SI, & SO	Off-the-shelf WiFi scanners, BT scanners	90% of observed unique IDs were WiFi addresses; popularity assessment experiments in 6 locations	-		Technology comparison for crowd monitoring
[?]*	SIM, MT, & MO	Custom sensors	Simulation for a train station; zoo experiment: 5 sensors, 7 days, 7K detected devices; sensors with 20m antennas indoors and 100m outdoors	-	Analytic method	Pedestrian quantity estimation
[?]*	ST, LI	Custom sensors	17 sensors, football stadium during a match; 15m range; 47,589 data points, 553 unique devices; average visited locations/device: 4.37, median: 2	14%	No	Sensor placement, pedestrian quantity estimation
[20]	LT, LI, & MO	Custom sensors	4 datasets: 7, 15, 17, 2 sensors; airport: 4 months, 141K addresses, 16.5M records; zoo: 14 days, 2K addresses, 150K records; F1 arena: 2 days, 12K addresses, 792K records; stadium: 8h, 2.5K addresses, 24K records	5%–12%	Location specific: turnstile data, video	Bluetooth's potential for acquiring pedestrian mobility data
[79]	ST, LI	Several types of Android phones	10 volunteers grouped in pairs, moving on a predefined path, some stationary, some walking; 4h of data, 3-day event, 1,600m^2 area, 1K+ visitors; various crowd densities; 75% classification accuracy	-	Video	Crowd density estimation
[59]	ST, MI	Laptop	3 days, 1 laptop placed at the store's entrance	30%	Manual observations, surveys	Acquiring shoppers' behavior using Bluetooth
[66]	LT, MI	Laptops	16 days, 2 laptops; WiFi and Bluetooth scans; 6,211 unique WiFi devices per day; 250 unique BT devices per day	-	Venue specific: airport gates data	Technology comparison for crowd sensing
[31]	MT, LO	Smartphone BLE tags	6day event, 2M visitors; 732 volunteers wearing tags; various demographics; 732 tags, 740K tag detections	98% (tags)	None	Event monitoring, group dynamics
[52]	LT, LO	Custom sensors	1 month, 14 sensors placed on billboards in Bonn; 5M data points, 85K devices; 7%–10% detection rate for another dataset	-	Similar dataset	Mobility patterns, commuter patterns
[54]	ST, LO	Phones	Festival, 280,000m^2; 10 agents, 13h; discovered 2,637 out of 3,326 BT devices; dynamic measurements	8.20%	Entrance scanners	Population size, density estimation
[58]	LT, LO	Off the shelf	3 weeks; 123 intersections with inductive loop detectors; 28 BT scanners placed on traffic lights/lamp posts; per scanner: 21 simultaneous connections, 3 antennas, 56-bit encryption	-	Inductive loop detectors	Multimodal traffic sensing

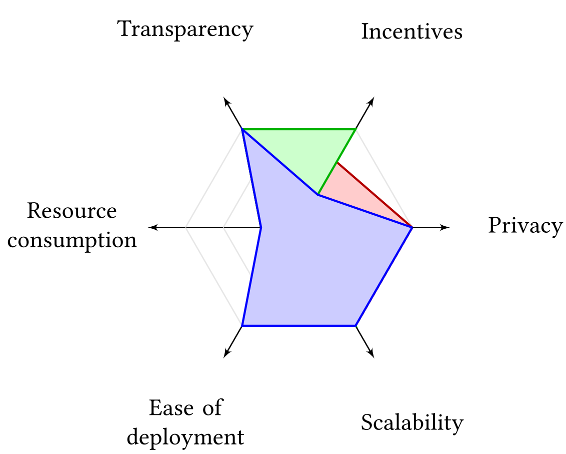
Table 8. Spot-On Papers

Solution	App Driven		Infrastructure Based, Hybrid		
	Mass Events	Urban Centric	Mass Events	Indoor Events	Urban Centric
[50]	-	✓	-	-	-
[70]	-	-	✓	-	-
[78]	-	-	✓	-	-
[33]	✓	-	-	-	-
[38]	-	-	-	✓	-
[53]	-	-	-	-	✓
[75]	-	-	✓	-	-
[1]	-	-	-	-	✓
[6]	-	-	✓	-	✓
[11]	-	-	✓	-	-
[14]	-	✓	-	-	-
[20]	-	-	✓	✓	-
[71]	✓	-	-	-	-
[79]	-	-	✓	✓	-
[83]	✓	-	-	-	-
[9]	✓	-	-	-	-
[13]	-	✓	-	-	-
[23]	-	-	-	-	✓
[56]	-	✓	-	-	-
[65]	-	-	-	✓	-
[66]	-	-	-	✓	-
[80]	✓	-	-	-	-
[31]	-	-	✓	-	-
[42]	-	-	-	-	✓
[48]	✓	-	-	-	-
[52]	-	-	-	-	✓
[54]	-	-	✓	-	-
[2]	-	-	✓	✓	-
[12]	-	-	✓	-	✓
[77]	-	-	✓	✓	-



		P	I	T	R	D	S
[33]	red	3	0	3	3	2	2
[73]	green	2	2	0	2	1	2
[82]	blue	3	1	3	2	2	3

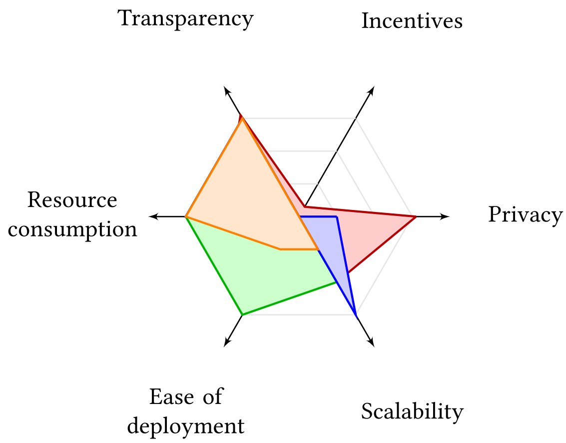
(a) Bluetooth, audio



		P	I	T	R	D	S
[85]	red	3	2	3	1	2	3
[9]	green	0	3	3	1	2	3
[50]	blue	3	1	3	1	3	3

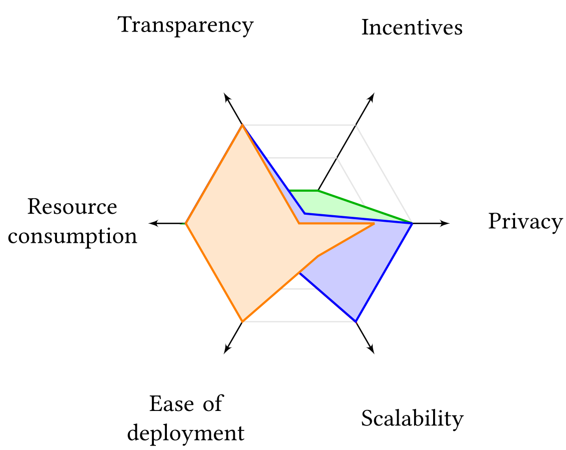
(b) GPS

Fig. 3. App-driven systems used in mass events.



		P	I	T	R	D	S
[72]		3	0	3	2	2	2
[80]		0	0	3	3	3	2
[77]		1	0	3	3	0	3
[56]		0	0	3	3	1	1

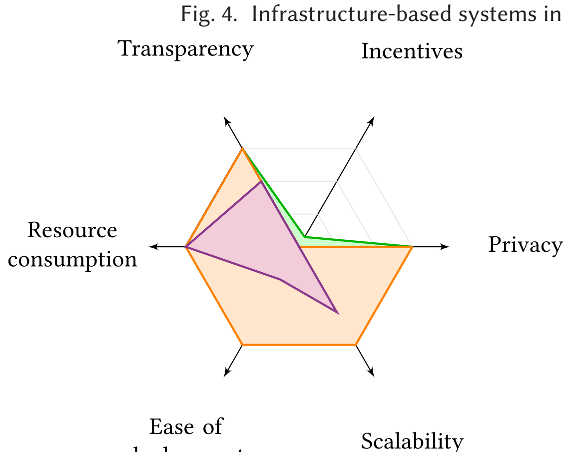
(a) Bluetooth



		P	I	T	R	D	S
[11]		1	0	3	3	1	3
[31]		3	1	1	3	0	1
[12]		3	0	3	3	1	3
[79]		2	0	3	3	1	3

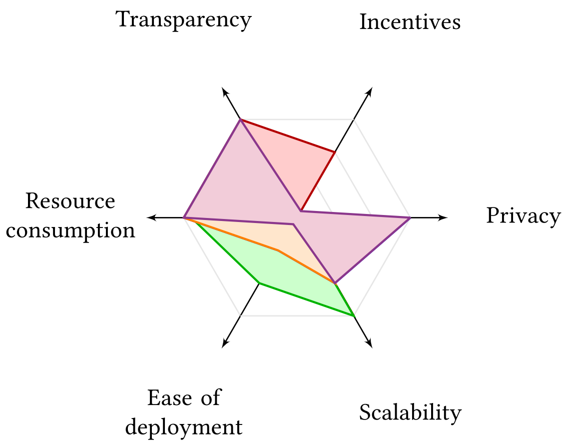
(b) Wi-Fi

Fig. 4. Infrastructure-based systems in mass events.



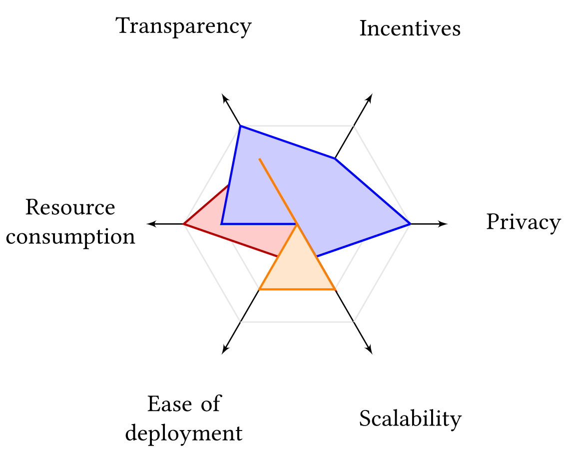
		P	I	T	R	D	S
[38]		0	0	3	2	2	0
[20]		3	0	3	3	1	2
[81]		0	0	3	3	2	1
[67]		3	0	3	3	3	3
[2]		0	0	2	3	1	2

Fig. 5. Infrastructure-based systems for indoor events.



		P	I	T	R	D	S
[55]		0	2	3	2	0	3
[6]		0	0	3	3	2	3
[23]		3	0	3	3	1	2
[42]		0	0	3	3	1	2
[54]		3	0	3	3	0	2

(a) Infrastructure-based



		P	I	T	R	D	S
[52]		3	1	1	2	2	3
[14]		3	2	2	2	0	1
[13]		3	2	3	2	0	1
[58]		0	0	2	0	2	2

(b) Application-driven

Fig. 6. Urban crowd-centric systems.

REFERENCES

[1] Naeim Abedi, Ashish Bhaskar, and Edward Chung. 2013. Bluetooth and WiFi MAC address-based crowd data collection and monitoring: Benefits, challenges and enhancement. In *36th Australian Transport Research Forum*.

[2] Utku Gunay Acer, Geert Vanderhulst, Afra Masshadi, Aidan Boran, Claudio Forlivesi, Philipp M. Scholl, and Fahim Kawsar. 2016. Capturing personal and crowd behavior with WiFi analytics. In *3rd International Workshop on Physical Analytics*. ACM Press, New York,43–48.

[3] Yutaka Arakawa and Yuki Matsuda. 2016. Gamification mechanism for enhancing participatory urban sensing: Survey and practical results. *Journal of Information Processing* 24, 1 (2016), 31–38.

[4] Paramvir Bahl and Venkata N. Padmanabhan. 2000. RADAR: An in-building RF-based user location and tracking system. In *19th INFOCOM Conference*, Vol. 2. IEEE Computer Society Press, Los Alamitos, CA, 775–784.

[5] Mehedi Bakht, John Carlson, Alexander Loeb, and Robin Kravets. 2012. United we find: Enabling mobile devices to cooperate for efficient neighbor discovery. In *12th Workshop on Mobile Computing Systems & Applications*. ACM Press, New York, 11.

[6] Marco V. Barbera, Alessandro Epasto, Alessandro Mei, Vasile C. Perta, and Julinda Stefa. 2013. Signals from the crowd: Uncovering social relationships through smartphone probes. In *13th Internet Measurement Conference*. ACM Press, New York, 265–276.

[7] Surya Bhattacharya, Henrik Blunck, Mikkel Baun Kjærgaard, and Petteri Nurmi. 2015. Robust and energy-efficient trajectory tracking for mobile devices. *IEEE Transations on Mobile Computing* 14, 2 (2015), 430–443.

[8] Ulf Blanke, Robin Guldener, Sebastian Feese, and Gerhard Tröster. 2014. Crowdsourced pedestrian map construction for short-term city-scale events. In *1st International Conference IoT in Urban Space*. ICST, 25–31.

[9] Ulf Blanke, Gerhard Troster, Tobias Franke, and Paul Lukowicz. 2014. Capturing crowd dynamics at large scale events using participatory gps-localization. In *9th International Conference Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP'14)*. IEEE Computer Society Press, Los Alamitos, CA, 1–7.

[10] Henrik Blunck, Niels Olof Bouvin, Tobias Franke, Kaj Grønbæk, Mikkel B. Kjaergaard, Paul Lukowicz, and Markus Wüstenberg. 2013. On heterogeneity in mobile sensing applications aiming at representative data collection. In *International Conference on Pervasive Computing and Communications (PerCom'13) (Adjunct Publications)*. ACM Press, New York, 1087–1098.

[11] Bram Bonné, Arno Barzan, Peter Quax, and Wim Lamotte. 2013. WiFiPi: Involuntary tracking of visitors at mass events. In *14th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM'13)*. IEEE, 1–6.

[12] Crisitian Chilipirea, Andreea Petre, Ciprian Dobre, and Maarten van Steen. 2016. Presumably simple: Monitoring crowds using WiFi. In *17th International Conference on Mobile Data Management*, Vol. 1. IEEE Computer Society Press, Los Alamitos, CA, 220–225.

[13] Yohan Chon, Suyeon Kim, Seungwoo Lee, Dongwon Kim, Yungeun Kim, and Hojung Cha. 2014. Sensing WiFi packets in the air: Practicality and implications in urban mobility monitoring. In *Joint Conference on Pervasive and Ubiquitous Computing (UbiComp'14)*. ACM Press, New York, 189–200.

[14] Yohan Chon, Nicholas D. Lane, Yunjong Kim, Feng Zhao, and Hojung Cha. 2013. Understanding the coverage and scalability of place-centric crowdsensing. In *Joint Conference on Pervasive and Ubiquitous Computing (UbiComp'13)*. ACM Press, New York, 3–12.

[15] Delphine Christin. 2016. Privacy in mobile participatory sensing: Current trends and future challenges. *Journal of Systems and Software* 116 (2016), 57–68.

[16] Delphine Christin, Andreas Reinhardt, Salil S. Kanhere, and Matthias Hollick. 2011. A survey on privacy in mobile participatory sensing applications. *Journal of Systems and Software* 84, 11 (2011), 1928–1946.

[17] Levent Demir, Mathieu Cunche, and Cédric Lauradoux. 2014. Analysing the privacy policies of Wi-Fi trackers. In *1st International Workshop on Physical Analytics*. ACM Press, New York, 39–44.

[18] Dustin Demuth, Daniel Nuest, Arne Bröring, and Edzer Pebesma. 2013. The airquality sensebox. In *EGU General Assembly Conference Abstracts*, Vol. 15. EGU2013–5146.

[19] Adriana Draghici. 2017. *Contributions to Mobility Data Management*. Ph.D. Dissertation. University Polytehnica Bucharest.

[20] Timothy Ellersiek, Gennady Andrienko, Natalia Andrienko, Dirk Hecker, Hendrik Stange, and Marc Mueller. 2013. Using Bluetooth to track mobility patterns: Depicting its potential based on various case studies. In *5th International Workshop on Indoor Spatial Awareness*. ACM Press, New York, 1–7.

[21] Frédéric Evennou and François Marx. 2006. Advanced integration of WiFi and inertial navigation systems for indoor mobile positioning. *EURASIP Journal on Advances in Signal Processing*1 (2006), 1–11.

[22] Julien Freudiger. 2015. How talkative is your mobile device?: An experimental study of Wi-Fi probe requests. In *8th Conference Security & Privacy in Wireless and Mobile Networks*. ACM Press, New York, 8:1–8:6.

[23] Yuki Fukuzaki, Masahiro Mochizuki, Kazuya Murao, and Nobuhiko Nishio. 2014. A pedestrian flow analysis system using Wi-Fi packet sensors to a real environment. In *Joint Conference on Pervasive and Ubiquitous Computing (UbiComp'14) (Adjunct Publications)*. ACM Press, New York, 721–730.

[24] Yuuki Fukuzaki, Masahiro Mochizuki, Kazuya Murao, and Nobuhiko Nishio. 2015. Statistical analysis of actual number of pedestrians for Wi-Fi packet-based pedestrian flow sensing. In *Joint Conference on Pervasive and Ubiquitous Computing (UbiComp'15)/International Symposium on Wearable Computing*. ACM Press, New York, 1519–1526.

[25] Raghu K. Ganti, Fan Ye, and Hui Lei. 2011. Mobile crowdsensing: Current state and future challenges. *IEEE Communications Magazine* 49, 11 (2011), 32–39.

[26] Bin Guo, Zhiwen Yu, Xingshe Zhou, and Daqing Zhang. 2014. From participatory sensing to mobile crowd sensing. In *International Conference on Pervasive Computing and Communications (PerCom'14) (Workshops)*. IEEE Computer Society Press, Los Alamitos, CA, 593–598.

[27] Bo Han and Aravind Srinivasan. 2012. eDiscovery: Energy efficient device discovery for mobile opportunistic communications. In *20th International Conference on Network Protocols*. IEEE Computer Society Press, Los Alamitos, CA, 1–10.

[28] Takamasa Higuchi, Hirozumi Yamaguchi, and Teruo Higashino. 2014. Context-supported local crowd mapping via collaborative sensing with mobile phones. *Pervasive and Mobile Computing* 13 (2014), 26–51.

[29] Takamasa Higuchi, Hirozumi Yamaguchi, and Teruo Higashino. 2015. Mobile devices as an infrastructure: A survey of opportunistic sensing technology. *Journal of Information Processing* 23, 2 (2015), 94–104.

[30] Hagen Höpfner and Maximilian Schirmer. 2012. Energy efficient continuous location determination for pedestrian information systems. In *11th International Workshop on Data Engineering for Wireless and Mobile Access.* ACM Press, New York, 58–65.

[31] Shuja Jamil, Anas Basalamah, Ahmed Lbath, and Moustafa Youssef. 2015. Hybrid participatory sensing for analyzing group dynamics in the largest annual religious gathering. In *Joint Conference on Pervasive and Ubiquitous Computing (UbiComp'15).* ACM Press, New York, 547–558.

[32] Eftychia Kalogianni, R. Sileryte, Marco Lam, Kaixuan Zhou, Martijn Van der Ham, S. C. Van der Spek, and E. Verbree. 2015. Passive WiFi monitoring of the rhythm of the campus. In *18th AGILE International Conference on Geographic Information Science.*

[33] Pravein Govindan Kannan, Seshadri Padmanabha Venkatagiri, Mun Choon Chan, Akhihebbal L. Ananda, and Li-Shiuan Peh. 2012. Low cost crowd counting using audio tones. In *10th Conference on Embedded Network Sensor Systems.* ACM Press, New York, 155–168.

[34] Mikkel Baun Kjærgaard and Henrik Blunck. 2014. Tool support for detection and analysis of following and leadership behavior of pedestrians from mobile sensing data. *Pervasive and Mobile Computing* 10 (2014), 104–117.

[35] Mikkel Baun Kjærgaard, Henrik Blunck, Markus Wüstenberg, Martin Wirz, Daniel Roggen, and Gerhard Tröster. 2013. Time-lag method for detecting following and leadership behavior of pedestrians from mobile sensing data. In *International Conference on Pervasive Computing and Communications (PerCom'13).* IEEE Computer Society Press, Los Alamitos, CA, 56–64.

[36] Mikkel Baun Kjærgaard and Marco Kuhrmann. 2015. On architectural qualities and tactics for mobile sensing. In *11th International Conference Quality of Software Architectures.* ACM Press, New York, 63–72.

[37] Mikkel Baun Kjærgaard, Jakob Langdal, Torben Godsk, and Thomas Toftkjær. 2009. Entracked: Energy-efficient robust position tracking for mobile devices. In *7th International Conference on Mobile Systems, Applications, and Services (MobiSys'09).* ACM Press, New York, 221–234.

[38] Mikkel Baun Kjærgaard, Martin Wirz, Daniel Roggen, and Gerhard Tröster. 2012. Detecting pedestrian flocks by fusion of multi-modal sensors in mobile phones. In *Joint Conference on Pervasive and Ubiquitous Computing (UbiComp'12).* ACM Press, New York, 240–249.

[39] Mikkel Baun Kjærgaard, Martin Wirz, Daniel Roggen, and Gerhard Tröster. 2012. Mobile sensing of pedestrian flocks in indoor environments using wifi signals. In *International Conference on Pervasive Computing and Communications (PerCom'12).* IEEE Computer Society Press, Los Alamitos, CA, 95–102.

[40] Robin Kravets, Hilfi Alkaff, Andrew Campbell, Karrie Karahalios, and Klara Nahrstedt. 2013. CrowdWatch: Enabling in-network crowd-sourcing. In *2nd ACM SIGCOMM Workshop on Mobile Cloud Computing.* ACM Press, New York, 57–62.

[41] Juong-sik Lee. 2010. Sell your experiences: A market mechanism based incentive for participatory sensing. In *International Conference on Pervasive Computing and Communications (PerCom'10).* IEEE Computer Society Press, Los Alamitos, CA, 60–68.

[42] Kai Li, Chau Yuen, and Salil Kanhere. 2015. SenseFlow: An experimental study of people tracking. In *6th Workshop on Real-World Wireless Sensor Networks.* ACM Press, New York, 31–34.

[43] Janne Lindqvist, Tuomas Aura, George Danezis, Teemu Koponen, Annu Myllyniemi, Jussi Mäki, and Michael Roe. 2009. Privacy-preserving 802.11 access-point discovery. In *2nd Conference on Wireless Network Security.* ACM Press, New York, 123–130.

[44] Markus Loecher and Tony Jebara. 2009. CitySense: Multiscale space time clustering of GPS points and trajectories. In *Proceedings of the Joint Statistical Meeting.*

[45] Elsa Macias, Alvaro Suarez, and Jaime Lloret. 2013. Mobile sensing systems. *Sensors* 13, 12 (2013), 17292–17321.

[46] Yuki Maekawa, Akira Uchiyama, Hirozumi Yamaguchi, and Teruo Higashino. 2014. Car-level congestion and position estimation for railway trips using mobile phones. In *Joint Conference on Pervasive and Ubiquitous Computing (UbiComp'14).* ACM Press, New York, 939–950.

[47] Marco Maier, Lorenz Schauer, and Florian Dorfmeister. 2015. ProbeTags: Privacy-preserving proximity detection using Wi-Fi management frames. In *11th International Conference on Wireless & Mobile Computing, Networking & Communication.* IEEE Computer Society Press, Los Alamitos, CA, 756–763.

[48] Joseph El Mallah, Francesco Carrino B, Omar Abou Khaled, and Elena Mugellini. 2015. Crowd monitoring critical situations prevention using smartphones and group detection. In *3rd International Conference on Distributed, Ambient and Pervasive Interactions (Lecture Notes in Computer Science)*, Vol. 9189. Springer-Verlag, Berlin, 496–505.

[49] Emiliano Miluzzo, Nicholas D. Lane, Kristóf Fodor, Ronald Peterson, Hong Lu, Mirco Musolesi, Shane B. Eisenman, Xiao Zheng, and Andrew T. Campbell. 2008. Sensing meets mobile social networks: The design, implementation and evaluation of the Cenceme application. In *6th Conference on Embedded Network Sensor Systems.* ACM Press, New York, 337–350.

[50] Emiliano Miluzzo, Michela Papandrea, Nicholas D. Lane, Andy M. Sarroff, Silvia Giordano, and Andrew T. Campbell. 2011. Tapping into the vibe of the city using Vibn, a continuous sensing application for smartphones. In *1st International Symposium on From Digital Footprints to Social and Community Intelligence*. ACM Press, New York, 13–18.

[51] Shinsuke Mori, Yu-Chih Wang, Takaaki Umedu, Akihito Hiromori, Hitoshi Yamaguchi, and Takeshi Higashino. 2012. Design and architecture of cloud-based mobile phone sensing middleware. In *2nd Symposium on Network Cloud Computing and Applications (NCCA'12)*. IEEE Computer Society Press, Los Alamitos, CA, 102–109.

[52] Marc Mueller, Daniel Schulz, Michael Mock, and Dirk Hecker. 2015. Detecting mobility patterns with stationary Bluetooth sensors: A real-world case study. In *18th AGILE International Conference on Geographic Information Science*.

[53] A. Musa and Jakob Eriksson. 2012. Tracking unmodified smartphones using wi-fi monitors. In *10th Conference on Embedded Network Sensor Systems*. ACM Press, New York, 281–294.

[54] Farid Movahedi Naini, Olivier Dousse, Patrick Thiran, and Martin Vetterli. 2015. Opportunistic sampling for joint population size and density estimation. *IEEE Transactions on Mobile Computing* 14, 12 (2015), 2530–2543.

[55] Tom Nicolai and Holger Kenn. 2007. About the relationship between people and discoverable Bluetooth devices in urban environments. In *4th International Conference Mobile Technology, Applications, and Systems*. ACM Press, New York, 72–78.

[56] Tomohiro Nishimura, Takamasa Higuchi, Hirozumi Yamaguchi, and Teruo Higashino. 2014. Detecting smoothness of pedestrian flows by participatory sensing with mobile phones. In *International Symposium on Wearable Computers*. ACM Press, New York, 15–18.

[57] Eamonn O'Neill, Vassilis Kostakos, Tim Kindberg, Alan Penn, Danaë Stanton Fraser, and Tim Jones. 2006. Instrumenting the city: Developing methods for observing and understanding the digital cityscape. In *Ubiquitous Computing (UbiComp'06)*. Lecture Notes in Computer Science, Vol. 4206. Springer-Verlag, Berlin, 315–332.

[58] Mikko Perttunen, Vassilis Kostakos, Jukka Riekki, and Timo Ojala. 2015. Urban traffic analysis through multi-modal sensing. *Personal and Ubiquitous Computing* 19, 3–4 (2015), 709–721.

[59] Peilin Phua, Bill Page, and Svetlana Bogomolova. 2014. Validating Bluetooth logging as metric for shopper behaviour studies. *Journal of Retailing and Consumer Services* 22 (2014), 158–163.

[60] Moo-ryong Ra, Bin Liu, Tom La Porta, and Ramesh Govindan. 2012. Medusa: A programming framework for crowd-sensing applications categories and subject descriptors. In *10th International Conference on Mobile Systems, Applications, and Services (MobiSys)*. Springer-Verlag, Berlin, 337–350.

[61] Kiran K. Rachuri, Christos Efstratiou, Ilias Leontiadis, Cecilia Mascolo, and Peter J. Rentfrow. 2013. METIS: Exploring mobile phone sensing offloading for efficiently supporting social sensing applications. In *International Conference on Pervasive Computing and Communications (PerCom'13)*. IEEE Computer Society Press, Los Alamitos, CA, 85–93.

[62] Francesco Restuccia, Sajal K. Das, and Jamie Payton. 2016. Incentive mechanisms for participatory sensing: Survey and research challenges. *ACM Transactions on Sensor Networks* 12, 2 (2016), 13.

[63] Daniel Roggen, Martin Wirz, Gerhard Tröster, and Dirk Helbing. 2011. Recognition of crowd behavior from mobile sensors with pattern analysis and graph clustering methods. *arXiv Preprint arXiv:1109.1664* (2011).

[64] Pierre Rouveyrol, Patrice Raveneau, and Mathieu Cunche. 2015. Large scale Wi-Fi tracking using a botnet of wireless routers. In *Workshop on Surveillance & Technology*.

[65] Antonio J. Ruiz-Ruiz, Henrik Blunck, Thor S. Prentow, Allan Stisen, and Mikkel B. Kjærgaard. 2014. Analysis methods for extracting knowledge from large-scale WiFi monitoring to inform building facility planning. In *International Conference on Pervasive Computing and Communications (PerCom'14)*. IEEE Computer Society Press, Los Alamitos, CA, 130–138.

[66] Lorenz Schauer, Martin Werner, and Philipp Marcus. 2014. Estimating crowd densities and pedestrian flows using WiFi and Bluetooth. In *11th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*. ACM Press, New York, 171–177.

[67] Katie Shilton. 2010. Participatory sensing: Building empowering surveillance. *Surveillance & Society* 8, 2 (2010), 131–150.

[68] Katie Shilton. 2012. Participatory personal data: An emerging research. *Journal of the Association for Information Science and Technology* 63, 10 (2012), 1905–1915.

[69] Katie Shilton, Jeffrey A. Burke, Deborah Estrin, Mark Hansen, and Mani Srivastava. 2008. Participatory privacy in urban sensing. In *International Workshop on Mobile Device and Urban Sensing*. IEEE Computer Society Press, Los Alamitos, CA.

[70] Hendrik Stange, Thomas Liebig, Dirk Hecker, Gennady Andrienko, and Natalia Andrienko. 2011. Analytical workflow of monitoring human mobility in big event settings using bluetooth. In *3rd International Workshop on Indoor Spatial Awareness*. ACM Press, New York, 51–58.

[71] Arkadiusz Stopczynski, Jakob Eg Larsen, Sune Lehmann, Lukasz Dynowski, and Marcos Fuentes. 2013. Participatory Bluetooth sensing: A method for acquiring spatio-temporal data about participant mobility and interactions at large

scale events. In *International Conference on Pervasive Computing and Communications (PerCom'13) Workshops*. IEEE Computer Society Press, Los Alamitos, CA, 242–247.

[72] Takumi Takafuji, Kazuhisa Fujita, Takamasa Higuchi, Akihito Hiromori, Hirozumi Yamaguchi, and Teruo Higashino. 2014. Indoor localization utilizing tracking scanners and motion sensors. In *11th International Conference Ubiquitous Intelligence and Computing*. IEEE Computer Society Press, Los Alamitos, CA, 112–119.

[73] Thiago Teixeira, Gershon Dublon, and Andreas Savvides. 2010. A survey of human-sensing: Methods for detecting presence, count, location, track, and identity. *Computer Surveys* 5 (2010), 1–77.

[74] Mathy Vanhoef, Célestin Matte, Mathieu Cunche, Leonardo S. Cardoso, and Frank Piessens. 2016. Why MAC address randomization is not enough: An analysis of Wi-Fi network discovery mechanisms. In *11th AsiaConference Computer and Communications Security*. ACM Press, New York, 413–424.

[75] Mathias Versichele, Tijs Neutens, Matthias Delafontaine, and Nico Van de Weghe. 2012. The use of Bluetooth for analysing spatiotemporal dynamics of human movement at mass events: A case study of the ghent festivities. *Applied Geography* 32, 2 (2012), 208–220.

[76] Yusuke Wada, Takamasa Higuchi, Hirozumi Yamaguchi, and Teruo Higashino. 2013. Accurate positioning of mobile phones in a crowd using laser range scanners. In *9th International Conference on Wireless & Mobile Computing, Networking & Communication*. IEEE Computer Society Press, Los Alamitos, CA, 430–435.

[77] Jens Weppner, Benjamin Bischke, and Paul Lukowicz. 2016. Monitoring crowd condition in public spaces by tracking mobile consumer devices with Wifi interface. In *Joint Conference on Pervasive and Ubiquitous Computing (UbiComp'16) (Adjunct Publications)*. ACM Press, New York, 1363–1371.

[78] Jens Weppner and Paul Lukowicz. 2011. Collaborative crowd density estimation with mobile phones. In *2nd International Workshop on Sensing Applications on Mobile Phones*. ACM Press, New York.

[79] Jens Weppner and Paul Lukowicz. 2013. Bluetooth based collaborative crowd density estimation with mobile phones. In *International Conference on Pervasive Computing and Communications (PerCom'13)*. IEEE Computer Society Press, Los Alamitos, CA, 193–200.

[80] Jens Weppner, Paul Lukowicz, Ulf Blanke, and G. Troster. 2014. Participatory Bluetooth scans serving as urban crowd probes. *IEEE Sensors Journal* 14, 12 (2014), 4196–4206.

[81] Nanda Wijermans, Claudine Conrado, Maarten van Steen, Claudio Martella, and Jie Li. 2016. A landscape of crowd-management support: An integrative approach. *Safety Science* 86 (2016), 142–164.

[82] Martin Wirz, Tobias Franke, Eve Mitleton-Kelly, Daniel Roggen, Paul Lukowicz, and Gerhard Tröster. 2013. Coenosense: A framework for real-time detection and visualization of collective behaviors in human crowds by tracking mobile devices. In *European Conference on Complex Systems*. Springer-Verlag, Berlin, 353–361.

[83] Martin Wirz, Tobias Franke, Daniel Roggen, Eve Mitleton-Kelly, Paul Lukowicz, and Gerhard Tröster. 2013. Probing crowd density through smartphones in city-scale mass gatherings. *EPJ Data Science* 2 (June 2013), 5.

[84] Martin Wirz, Daniel Roggen, and Gerhard Troster. 2009. Decentralized detection of group formations from wearable acceleration sensors. In *International Conference Computational Science and Engineering*, Vol. 4. IEEE Computer Society Press, Los Alamitos, CA, 952–959.

[85] Martin Wirz, Pablo Schläpfer, Mikkel Baun Kjærgaard, Daniel Roggen, Sebastian Feese, and Gerhard Tröster. 2011. Towards an online detection of pedestrian flocks in urban canyons by smoothed spatio-temporal clustering of GPS trajectories. In *3rd International Workshop on Location-Based Social Networks*. ACM Press, New York, 17–24.

[86] Yu Xiao, Pieter Simoens, Padmanabhan Pillai, Kiryong Ha, and Mahadev Satyanarayanan. 2013. Lowering the barriers to large-scale mobile crowdsensing. In *14th Workshop on Mobile Computing Systems and Applications*. ACM Press, New York, 9.

[87] Huijing Zhao and Ryosuke Shibasaki. 2005. A novel system for tracking pedestrians using multiple single-row laser-range scanners. *IEEE Transactions on Systems, Man, and Cybernetics* 35, 2 (2005), 283–291.

Received October 2016; revised April 2017; accepted July 2017