REGULAR ARTICLE

Probing crowd density through smartphones in city-scale mass gatherings

Martin Wirz[1*], Tobias Franke[2], Daniel Roggen[1], Eve Mitleton-Kelly[3], Paul Lukowicz[2] and Gerhard Tröster[1]

*Correspondence:
martin.wirz@ife.ee.ethz.ch
[1] Wearable Computing Laboratory,
ETH Zürich, Zürich, Switzerland
Full list of author information is
available at the end of the article

Abstract

City-scale mass gatherings attract hundreds of thousands of pedestrians. These pedestrians need to be monitored constantly to detect critical crowd situations at an early stage and to mitigate the risk that situations evolve towards dangerous incidents. Hereby, the crowd density is an important characteristic to assess the criticality of crowd situations.

In this work, we consider location-aware smartphones for monitoring crowds during mass gatherings as an alternative to established video-based solutions. We follow a participatory sensing approach in which pedestrians share their locations on a voluntary basis. As participation is voluntarily, we can assume that only a fraction of all pedestrians shares location information. This raises a challenge when concluding about the crowd density. We present a methodology to infer the crowd density even if only a limited set of pedestrians share their locations. Our methodology is based on the assumption that the walking speed of pedestrians depends on the crowd density. By modeling this behavior, we can infer a crowd density estimation.

We evaluate our methodology with a real-world data set collected during the Lord Mayor's Show 2011 in London. This festival attracts around half a million spectators and we obtained the locations of 828 pedestrians. With this data set, we first verify that the walking speed of pedestrians depends on the crowd density. In particular, we identify a crowd density-dependent upper limit speed with which pedestrians move through urban spaces. We then evaluate the accuracy of our methodology by comparing our crowd density estimates to ground truth information obtained from video cameras used by the authorities. We achieve an average calibration error of 0.36 m^{-2} and confirm the appropriateness of our model. With a discussion of the limitations of our methodology, we identify the area of application and conclude that smartphones are a promising tool for crowd monitoring.

Keywords: crowd sensing; pedestrian behavior; crowd density estimation; participatory sensing; smartphone

1 Introduction

City-scale mass gatherings attract hundreds of thousands of attendees. On 25 April 2011, an estimated number of 1.2 million spectators congregated in London for the wedding of Prince William and Catherine Middleton [1]. Around 2 million people gathered on 25 May 2010 in Buenos Aires to attend several concerts and street art parades celebrating the Bicentennial of the May Revolution [2]. Up to 2 million people got together in Madrid, Spain for a parade celebrating the success of the Spanish national football team winning the 2010 FIFA World Cup [3]. Such events with many visitors but with a restricted area and com-

plex architectural configurations like narrowings and intersections bear the risk of dangerous crowd incidents [4, 5]. It is therefore a top priority for organizers of such events to maintain a high level of safety and to minimize the risk of crowd incidents. Hereby, guidelines on planning help minimize the risk by deploying adequate safety measures [6, 7]. The raise of pedestrian simulation tools has enabled the identification of critical locations where dangerous crowd behaviors may emerge [8, 9]. Simulation tools help to design and proactively deploy crowd control mechanisms before mass gatherings to mitigate the risk of dangerous crowd incidents. However, despite a proper preparation, the behavior of the crowd during an event remains highly unpredictable [10, 11]. Hence, emerging critical crowd situations need to be detected at an early stage in order to mitigate the risk of a situation evolving towards a dangerous incident. Crowd density, *i.e.* the number of people per unit area, has been identified as one important measure to assess the criticality of a situation [12, 13] and there is a need to obtain this information during an event [14].

In our ongoing research effort, we want to turn pedestrians' smartphones into a reliable sensing tool for measuring the crowd density during city-wide mass gatherings. In a previous study [14], we introduced a participatory sensing system for crowd monitoring by tracking the location of attendees of mass gatherings via their smartphones. Attendees of such a mass gathering can download a smartphone App to record the user's location at regular intervals. This information is collected from all App users and used to infer the users' current spatial distribution. To motivate as many attendees as possible to download the App and share their locations, the App offers a set of features including an interactive festival program and maps of the venue as an incentive to all. Nevertheless, by following a participatory sensing approach, we expect only a fraction of all attendees to participate and hence, the location of only a limited set of pedestrians is known. Therefore, the explanatory power of the obtained distribution is limited as these numbers do not provide direct evidence of the actual crowd density.

In this work, we address this challenge and present a methodology to infer the crowd density by tracking the locations of a subset of all event attendees. Our methodology relies on a calibration approach that provides a relation between the distribution of App users and the crowd density. Hereby, we make use of the characteristic that pedestrians exhibit a distinct behavior which depends on the crowd density in the vicinity. By assessing the behavior of the App users and applying our model, we obtain a crowd density estimation. Evaluation of our approach is performed with a real-world data set collected during the Lord Mayor's Show 2011 in London, a festival attracting around half a million spectators. We use this data set to confirm the suitability of our methodology and evaluate the accuracy of our crowd density estimation by comparing our results to results from video footage obtained from CCTV cameras. We conclude our work by addressing the limitations of our methodology and identifying next steps.

2 Related work

This section discusses related work. Section 2.1 introduces crowd characteristics relevant to assess the criticality of a situation during mass gatherings. Section 2.2 compares technologies and methods to measure such crowd characteristics with a focus on crowd density.

Table 1 Chart of crowd density

Dynamics	Density [m^{-2}]	Behavior and risk
Standing	7.1	Critical crowd density for static crowds
Walking	0.43	Stream of pedestrians can maintain normal walking speed and avoid one another
	2	Walking speed is reduced
	3.57	Involuntary contact is experienced between people
	5.55	Potentially dangerous crowd forces begin to develop

Chart derived from the findings of Fruin [23] to assess the criticality of a crowd density. The density of a static crowd can exceed the density of a moving crowd before a critical values is reached.

2.1 Crowd characteristics to assess the criticality of a situation during mass gatherings

Various empirical studies have analyzed crowd behaviors during mass gatherings and identified critical, potentially dangerous situations: A focus in literature has been the investigation of human stampedes [15–19]. Stampedes often occur if people start to rush towards a common target. Congestions, or clogging, at narrowings and counter flow of pedestrians have been identified as critical situations in which stampedes may occur [20, 21]. Irregular pedestrian flow is an additional risk which may cause turbulent motions in a crowd [20]. *Johansson et al.* [22] identified the transition from smooth pedestrian flow to stop-and-go waves as a warning sign of a critical situation.

Based on such observations, researchers have identified different crowd characteristics that may indicate potentially critical situations. One of the most important crowd characteristic is the local crowd density. *Au et al.* [13] report that one of the key aspects in developing and maintaining a crowd safety system is to identify areas where crowds build up. Areas where people are likely to congregate need careful observation during an event to provide crowd safety. *Nicholson et al.* [12] state the need for accurate crowd density estimation to correctly asses the criticality of a situation. Crowd density is also observed by police forces during the management of mass gatherings. Table 1 shows a chart derived from the findings of Fruin [23] to assess the criticality of a situation of a situation during a mass gathering.

The local crowd density alone does not allow for a complete assessment of the criticality of a situation. In addition to crowd density, the intention or behavior of a crowd is required for a correct situational understanding. As an example, a high crowd density in a static crowd is less critical than a high crowd density exhibiting counter flow. This distinction is also evident in Table 1. A critical crowd density is reached at 5.55 m^{-2} for a moving crowd. A static crowd, however, can exceed this value before a critical density is reached. *Helbing et al.* [20] introduce a measure that incorporates this aspect. They call this measure *crowd pressure* which is given as the local velocity variance multiplied by the local crowd density. In their work, they identified that crowd pressure can be seen as an early warning sign for critical crowd situations. They identified an increased crowd pressure value right before dangerous crowd turbulence emerges.

2.2 Monitoring crowds

Nowadays, video-based crowd monitoring tools are widely deployed. *Gong et al.* [24] review the state-of-the-art of vision-based systems for crowd monitoring. They conclude that currently deployed systems suffer from poor scalability to crowded public spaces due to deployment complexity and manually judging the criticality of a situation from

the footage. Further, manually monitoring multiple video streams simultaneously requires lots of training for a person. To overcome these limitations, police forces use helicopters to gain an instantaneous overview and men in the field to obtain detailed information [14].

Recent developments such as multi-camera networks to fuse information from multiple cameras and computer vision algorithms to automatically monitor crowds can mitigate these issues. *Jacques et al.* [25] review state-of-the-art techniques. Hereby, the authors differ between *object-based approaches* and *holistic approaches*. In object-based approaches, single individuals are detected and tracked individually. Relevant information is fused to analyze group behaviors. As an example, *Mehran et al.* [26] use the social-force model introduced by *Helbing et al.* [27] to infer crowd patterns from pedestrian tracks. Object-based approaches have been used by *Johansson et al.* [22] investigate crowd behaviors during the Hajj in Makkah. *Steffen et al.* [28] presented approaches for inferring crowd densities and other crowd behaviors based on pedestrian trajectories.

Holistic approaches do not rely on tracking individuals but follow a top-down methodology in which the crowd is considered as a single entity. These approaches obtain coarser-level information such as crowd density, the flow of the crowd and crowd turbulence but no local, individual-specific information. As an example, *Krausz et al.* [19] developed an optical flow-based method for an automatic detection of dangerous motion behaviors including congestions during mass gatherings. They used their method to study video-footage recorded during the Love Parade disaster of 2010 in Duisburg, Germany where 21 visitors died in a stampede. By comparing the two approaches, the authors of [25] write that while object-level analysis tends to produce more accurate results, the identification of individuals is challenging in high density crowds due to clutter and occlusion which makes it difficult to obtain an accurate estimation of the crowd density.

Despite the recent advances of computer vision and pattern recognition techniques, until now, it remains challenging to obtain an automated global situation awareness during mass gatherings from video footage [24]. Using alternative technologies for observing crowds has recently found interest in the research community. Hereby, thanks to their proliferation, mobile devices like smartphones have increasingly been considered as a viable tool for monitoring the behavior of a crowd. These sensor-rich devices offer various ways to obtain information about the whereabouts of their users and hence allow for monitoring the physical behavior of them [29]. By combining information from many people, the behavior of a collective can be monitored.

To infer crowd conditions like those mentioned in Section 2.1, the location of attendees of a mass gathering is required. There are different approaches to determine a smartphone's location which can broadly be divided into two classes: *in-network localization* and *on-device localization*. The in-network location methods utilize the fact that at any given time, a smartphone is connected to a cell tower in a network. The information which device is connected to which cell tower is being stored centrally in a database and updated constantly. Since the location of each cell tower is known, a position estimation of the mobile devices can be obtained. For on-device localization methods, on the other hand, the location is derived directly on the users' smartphones by means of GPS positioning, WiFi-fingerprinting or other comparable approaches [30]. The in-network localization approaches have the advantage that the locations of all subscribed devices are routinely being logged by the network operators. Thus, location information from a large number of devices can be obtained without any user interaction (and permission). Popular meth-

ods for obtaining in-network location estimation include the recording of network bandwidth usage by detecting how much communication is going on in a particular location. *Calabrese et al.* [31] used this measure to investigate crowd dynamics in the city of Rome. The obtained measure is an aggregated number which is highly dependent on communication behavior and is not necessarily correlated to the actual number of individuals in that location. Another method to capture in-network location information is to use Call Data Records (CDRs) [32, 33]. A single CDR tuple is generated for every voice call and Short Message Service (SMS) transaction and consists of the sender and receiver numbers together with a timestamp and the cell ID the sender is situated in. This data is routinely being collected by every network operator for operational and billing purposes. While being useful for many studies, CDR-based location data faces several limitations. Firstly, CDRs are sparse in time because they are generated only when a transaction occurs and not at fixed periodic intervals. Hence, as long as no communication takes place, a smartphone's location is not being revealed. Secondly, they are coarse in space as they record locations at the granularity of a cell tower sector resulting in a location uncertainty of around 300 meters [34].

Methods to obtain on-device location information include GPS positioning and WiFi/GSM-fingerprinting [35]. With these approaches a location accuracy of up to 5 m can be obtained for GPS and around 20 m for WiFi-based positioning, respectively [36, 37]. A further advantage is that in contrast to in-network methods, location updates of a user can be recorded at regular intervals and not sporadically, event-driven as in the case of CDRs. This makes it much simpler to extract movement trajectories and is less situational-biased as opposed to if positions are only recorded if communication is going on. *Koshak et al.* [38] use GPS positioning to track pedestrian movements in a crowded area in Makkah. With a post-event evaluation, they identified critical zones by evaluating the crowd flow obtained from the collected GPS updates. There are other means to track the location of smartphone users and estimate a crowd density. As an example, *Versichele et al.* [39] present an approach where Bluetooth beacons are placed in the environment in order to track smartphone users during a city-wide festival. The authors conduct a post-event evaluation to understand the spatial commuting pattern of the festival visitors. While Bluetooth can provide a fine-grained position estimation, it requires beacons placed in the environment to observe pedestrians and hence, people are only tracked at specific locations around deployed beacons. The work of Bandini discusses in [40] opportunities and challenges of different technologies for tracking pedestrians in crowded situations. Table 2 summarizes our literature review by listing different technologies and methods the assessment of the crowd density.

We conclude that determining the location of a person on a mobile device using GPS or any other localization approach can provide a much more accurate location estimation compared to in-network approaches. On-device localization methods also have advantages over vision-based approaches as limitations such as occlusion or the limitations in low-light conditions are inexistent and that the whole venue space can easily be covered. However, on-device localization approaches face a big challenge: In contrast to in-network methods, the location is determined on a user's smartphone. To collect this information, a user has to deliberately share it. This requires a dedicated piece software running on the device.

Table 2 Overview of technologies and methods for crowd density assessment

Sensor modality	Class	Method	Reference
Video	Holistic	Optical flow	[19]
		Fractal dimension	[41]
		Pixel counting	[42]
		Machine learning	[43]
	Object-based	Trajectories	[22, 28]
		Supervised classification	[44]
		Unsupervised clustering	[45]
		Social force model	[26]
Smartphone	In-network	CDR	[32, 33]
		Network bandwidth usage	[31, 46]
	On-device	Trajectories	[14, 38]
	Other	Bluetooth beacons	[39, 47]

Overview of technologies and methods to automatically to assess crowd density. Different video-based approaches have been investigated, a selection of methods is given here. Thanks to the location-awareness of modern smartphones, they have increasingly been considered as an alternative platform for crowd monitoring.

We present in the next section methods to infer crowd characteristics from location information as provided by smartphones. Afterwards, in Section 3, we will address the implications on-device localization approaches face by requiring people to run a piece of software on the smartphones. We then present our method to mitigate the influence.

2.3 Measures of local crowd characteristics and their relation

2.3.1 Crowd density and speed of the crowd
The density and speed of a crowd are important local characteristics to assess the criticality of a crowd situation. In this section, we present methods to derive these measures from position information of pedestrians and discuss their relation.

Local crowd density *Johansson et al.* [22] introduce the notation of local density $\rho(\vec{r}, t)$. The local density is determined by considering the location $\vec{r}_i$ of all pedestrians i at time t and is given as:

$$\rho(\vec{r}, t) = \frac{1}{\pi R^2} \sum_i \exp\left[-\left\|\vec{r}_i(t) - \vec{r}\right\|^2 / R^2\right], \tag{1}$$

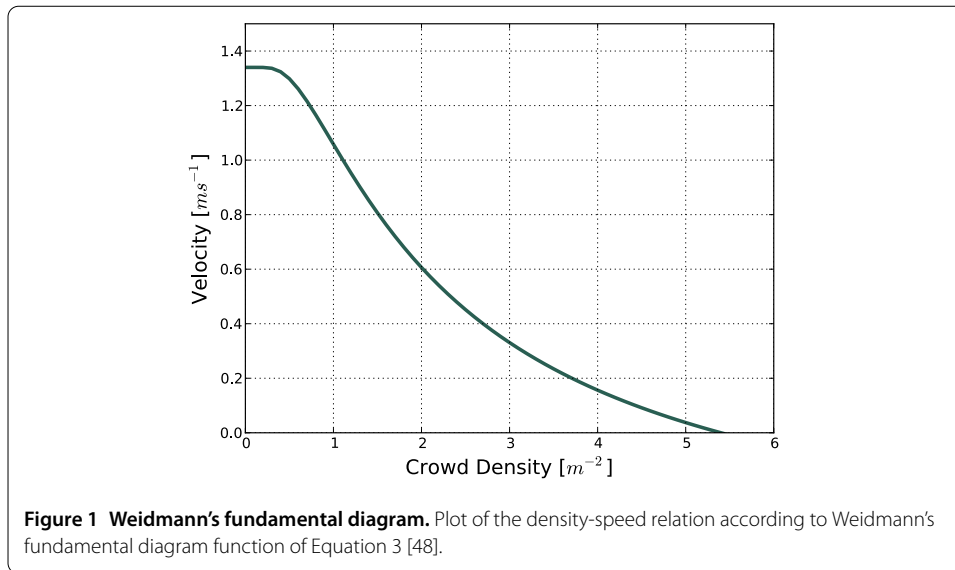where R is the kernel radius and defines the smoothing around the location $\vec{r}$.

Local crowd speed The local crowd speed is calculated in an analogous fashion as the crowd density [22]. To obtain a crowd speed value v, a weighted mean function is applied on the speed measures of the pedestrians around the location $\vec{r}$. Hence, the local speed is given as

$$v(\vec{r}, t) = \frac{\sum_i v_i \exp[-\|\vec{r}_i(t) - \vec{r}\|^2 / R^2]}{\sum_i \exp[-\|\vec{r}_i(t) - \vec{r}\|^2 / R^2]}, \tag{2}$$

where v_i is the speed of pedestrian i at location $\vec{r}_i$ and time t. Again, R is the kernel radius.

2.3.2 The fundamental diagram: relation between crowd density and speed
The influence of the crowd density on the walking speed of pedestrians has been investigated intensively for the purpose of dimensioning pedestrian facilities with respect to

Figure 1 Weidmann's fundamental diagram. Plot of the density-speed relation according to Weidmann's fundamental diagram function of Equation 3 [48].

comfort and safety. For low crowd density situations, pedestrians will be able to maintain free flow speed and are not interrupted by their neighbors. However with increasing density, the speed will decrease as the influence of the neighboring pedestrians force speed adjustments. This is similar to the situations in vehicular traffic [49]. This speed-density relationship is termed *Fundamental Diagram.* Weidmann [48] was one of the first to look at this relationship for pedestrians and proposed an analytical description from empirical data. He proposed to describe the relation between local density and speed as follows:

$$v(\rho) = v_0 \left\{ 1 - \exp\left[-\gamma \left(\frac{1}{\rho} - \frac{1}{\rho_{\max}} \right) \right] \right\}, \tag{3}$$

where $v_0 = 1.34$ ms^{-1} is the free speed at low densities (free flow), $\rho_{\max} = 5.4$ m^{-2} the maximal pedestrian density from which onward movement is not possible anymore and $\gamma = 1.913$ m^{-2} a fit parameter. Figure 1 shows a plot of the fundamental diagram given by Equation 3 and the listed parameters. The work of Weidmann stimulated successive contributions focusing on verifying and understanding this relationship. Several reports focus on the influence of various architectural configurations [50, 51], different crowd patterns [52] as well as demographics and cultural aspects [53, 54] on the fundamental diagram. Other works use the fundamental diagram to model pedestrian behaviors [55–57], investigate microscopic behavior patterns [58] and discuss and compare variations found across fundamental diagrams from different works [55, 59]. By comparing the results with other empirical data sets, it was found that the fundamental diagram is highly cultural dependent and needs to be adjusted for different venues. Weidmann's equation relies on fitting the fundamental diagram's analytical function to the recorded data set. Johansson addresses this issue in [55] and presents a generalized model. It relies on measurable parameters only and not on arbitrary fit parameters. Johannson showed that the model fits for different data sets. It can be tuned to follow existing models derived from various empirical data sets. Hence, the methods is believed to be sufficiently generic to be applied to various real-life situations. Johannson's method only relies on the maximum local crowd density and the free speed of pedestrians in unrestricted conditions. Both

parameters are highly cultural and demographic specific and hence are expected to vary significantly for different events. Nevertheless, the parameters are measurable and can be determined based on values from literature, expert knowledge or empirical measurements [53, 60].

3 Considering App users as probes to infer crowd characteristics
3.1 Challenges in participatory sensing systems

Section 2.2 discusses the advantages of on-device localization methods for tracking pedestrians and identifies a major challenge: In contrary to in-network approaches, people have to deliberately share their position information. This requires a dedicated piece of software running on a user's smartphone. At first sight, such an approach may appear undesirable, as it can be assumed that the majority of people is not willing to install such an application and constantly send their current position to a remote server for various reasons, including privacy concerns and energy considerations. In the case of a mass gathering, this may imply that only a fraction of all attendees would run such an application and many would opt for not having their location tracked. However, in a preceding study, we verified that people are willing to share privacy-sensitive location information if they receive some benefits or if they realize that sharing such information is for their own good and safety [61]. Thus, we believe such an approach is still viable and promising by following a participatory sensing scheme where users are motivated to deliberately share their location information by providing them with incentives and making it very transparent what the data is being used for. In [14] we introduce the concept of a smartphone App that tracks pedestrian's movements and offers attendees of a mass gathering a set of features which users regard as useful to them, *e.g.* an interactive program guide, a map superimposing the location of points of interest, or background information about the mass gathering. During the event, users of the App can receive location-dependent messages from the police. Through the users' smartphones, the police can inform users situated in a particular area with targeted information on how to behave in case of an emergency.

3.2 Considering App users as probes

Even by deploying an attractive App to reach a large user base, we can only expect to receive position information from a fraction of all event attendees. Our concept to infer crowd conditions by only tracking a limited number of event attendees is to consider the App users who share data as so called *probes* and extrapolate crowd information based on their behaviors. This is comparable to approaches in zoology where scientists monitor schools of fish or packs of mammals by equipping some of the members with tracking sensors to monitor and study interaction patterns and conclude about the whole group's social behavior and habitats. Following such an approach imposes a set of assumptions which we will discuss in the following:

1. Unknown ratio of App users: The ratio of event attendees using the App at any given moment is unknown. While the absolute number of App users is known, it is usually not possible to obtain the exact number of event attendees at a certain point in time.
2. Spatial distribution of App users corresponds to the distribution of event: Throughout the whole event we consider a spatial distribution of App users that corresponds to the spatial distribution of event attendees. This means that among

the event attendees, the App users are equally distributed. This is important, as it helps us to discover trends. While it does not allow us to directly infer how many people resist at one location, we can identify that a certain percentage of users, and hence event attendees, situates themselves in a given area.

3. Natural behaviors and interaction patterns: App users behave naturally and interact with the environment and other persons in a similar way as non-App-users. Hence, the averaged behavior of the App users at one specific location corresponds to the averaged behavior of the event attendees in this area. By accepting this assumption, we can infer certain crowd characteristics at a given location even if not every person is being tracked. We simply infer the behavior by considering the behavior of the App users. This is possible because pedestrians in crowds are likely to mimic the behavior of the neighboring pedestrians, *e.g.* by adjusting their walking speed and direction [62, 63]. By looking at a single individual, this assumption may not hold as a person may always decide independently on their behavior, *e.g.* stand still, walk in another direction, *etc.* However, by averaging over the App users, we assume that the averaged App user behavior corresponds to that of the crowd at a given location.

The more pedestrians participate and share their location, the more reliable we can conclude about occurring crowd characteristics. However, the obtained App user distribution does not reflect the actual crowd density. In the following section, we briefly cover the data collection platform and present the data set used for evaluation. Afterwards, we verify the assumptions introduced in this section and focus on the density-speed relation in our data set. Based on the obtained findings, in Section 5.6 we present our methodology to automatically infer a crowd density estimation from the collected position data and evaluate it against ground truth information obtained from video footage.

4 Data collection framework and data set

4.1 CoenoSense data collection framework

To collect location updates from pedestrians, we developed a generic App for mobile devices which can be tailored to a specific mass gathering and provides the users with event-related information and features. These features are designed to be attractive and useful during the event to reach a large user base. While a user's smartphone is running the App, the current location of the device is sampled at 1 Hz using the integrated GPS sensor. Such a high sampling rate was chosen to capture as much of the motion dynamics as possible. Besides the user's current location, the recorded GPS information also reveals the current velocity and heading direction of a user. This information is logged too. The recorded data is periodically sent a server running the *CoenoSense* framework. *CoenoSense* is a data collection backend infrastructure to collect and store arbitrary context information received from potentially thousands of mobile devices simultaneously. It allows for real-time processing of the collected data.

To ensure a user's privacy, data is sent anonymously and our App offers users full control over data sharing and data recording. It can be disabled by the user at any time.

4.2 Data set

We deployed the App and the *CoenoSense* platform during the Lord Mayor's Show 2011 which took place in London on November the 12th between 11 am and 6 pm. The Lord Mayor's Show is a street parade in the City of London, the historic core of London and

the present financial centre. The App offers a festival program, a map indicating points of interest and additional background information about the event. In collaboration with the event organizers, we event's official iPhone App and distributed it for free. It was advertised on the Lord Mayor's Show website and available through Apple's iTunes App store.

GPS location updates were collected between 00:01 on November 12th and 23:59 the same day and only if a user was in a specific geographical area around the venue the event takes place.
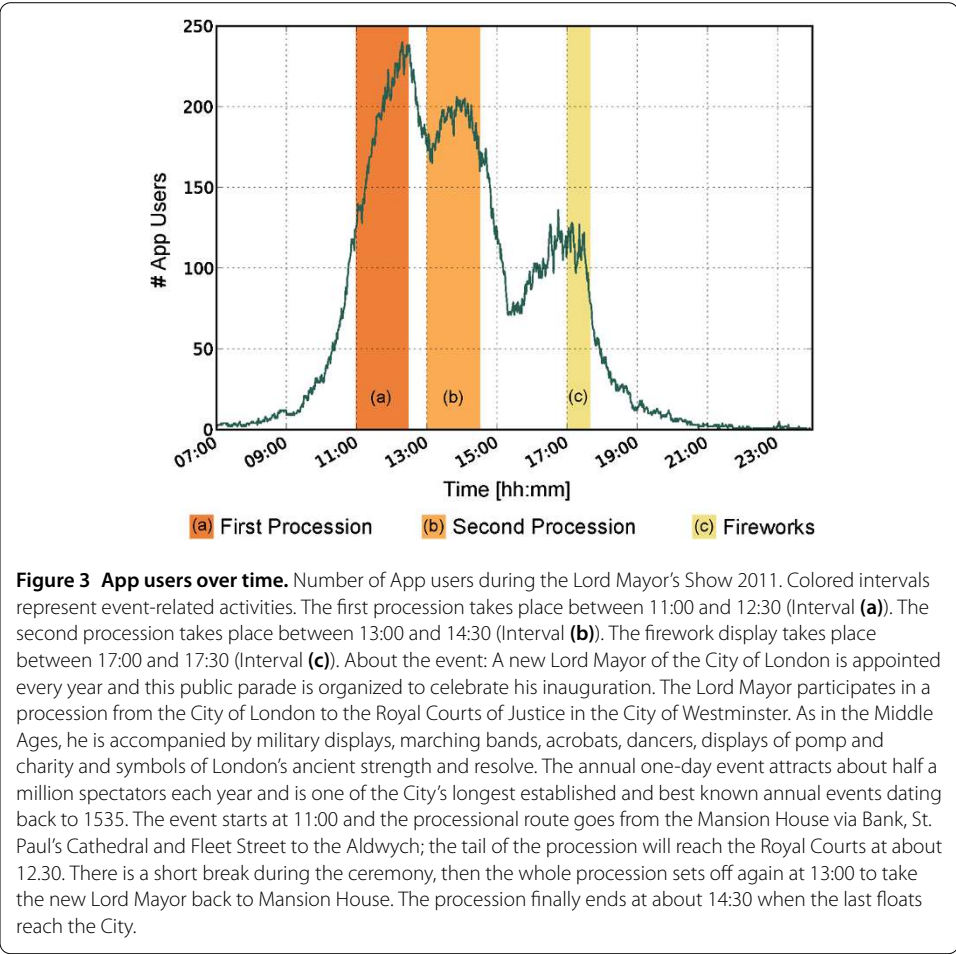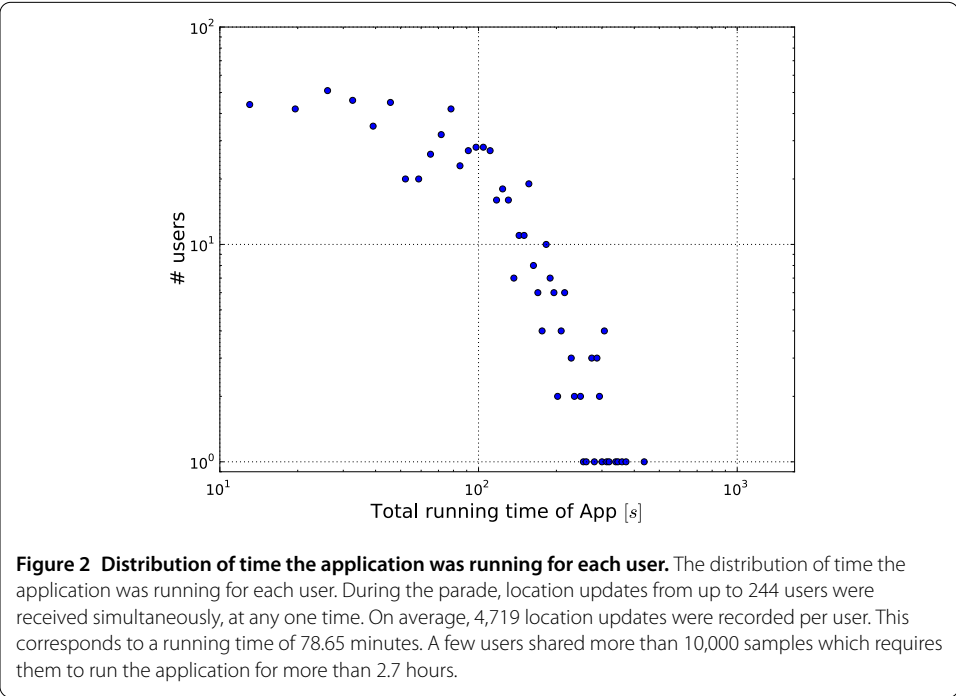
Within the collaboration with the event organizers and police forces, we obtained access to the CCTV video footage recorded during the Lord Mayor's Show. These are the same video recordings as used by the police to monitor the event. We consider this footage as ground truth information and is used in the following sections to verify our assumptions and evaluate our methods. We used video footage from four cameras placed at different locations. These locations have been identified by the police as being critical with respect to occurring crowd behaviors. For each camera, we defined an area of approximately 10 m^2 within which the crowd density is being extracted.
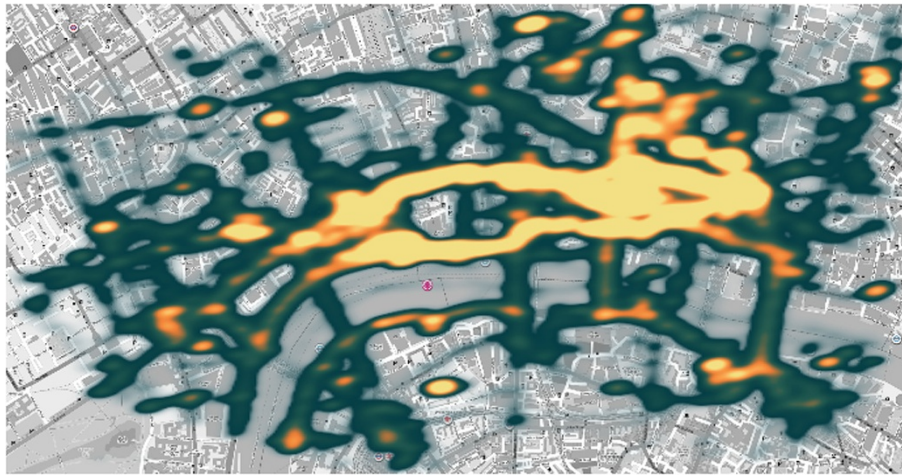
5 Empirical findings

In this section, we report on various spatio-temporal behavior properties that can be discovered in our data set. We start by investigating general statistics and put a special focus on aspects which help to support the assumptions stated in Section 3.2. Afterwards, we focus on the density-velocity relation.
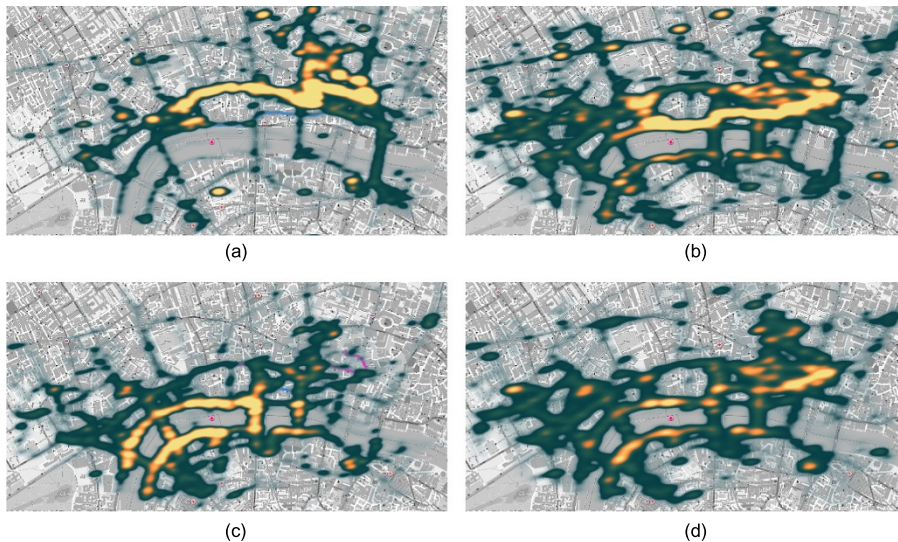
5.1 Spatio-temporal distribution of App users

We collected a total of 3,903,425 location updates from 828 different users. During the parade, location updates from up to 244 users were received simultaneously, at any one time. On average, 4,719 location updates were recorded per user. This corresponds to a running time of 78.65 minutes. A few users shared more than 10,000 samples which requires them to run the application for more than 2.7 hours. Figure 2 shows this by illustrating the distribution of time the application was running for each user. To understand the temporal usage pattern, Figure 3 shows the number of active users throughout the event. The axis of abscissae represents the time of the day. The axis of ordinate indicates the number of active users that share location updates at each point in time. Periods in which important event-related activities took place are indicated with a colored background. The first procession happens between 11:00 and 12:30 (Interval (a)). After a break, the second procession takes place between 13:00 and 14:30 (Interval (b)). Before the end of the event, a firework display takes place between 17:00 and about 17:30 (Interval (c)). Figure 4 shows the spatial usage pattern. Superimposed is a heat map representation of the spatial distribution of the collected data samples throughout the whole event. The heatmap visualizes the density of the reported location updates. The more data has been collected at a location, the 'hotter' (*i.e.* more yellow) it is colored. From this plot we can deduce that data collection is not uniform across space but concentrated to specific areas. These areas correspond to the locations in which event-related activities took place. However, in this plot, temporal information is lost. It does not allow to distinguish whether there is a high concentration of pedestrians for a short time or a few users stationary for a long time. To better understand spatio-temporal dynamics, Figure 5 shows the heat maps of four different time intervals. Hereby, Figure 5(a) shows the distribution of reported locations during the first procession (Interval (a)), Figure 5(b) during the second procession (Interval (b)) and Figure 5(c)

Figure 2 Distribution of time the application was running for each user. The distribution of time the application was running for each user. During the parade, location updates from up to 244 users were received simultaneously, at any one time. On average, 4,719 location updates were recorded per user. This corresponds to a running time of 78.65 minutes. A few users shared more than 10,000 samples which requires them to run the application for more than 2.7 hours.

Figure 3 App users over time. Number of App users during the Lord Mayor's Show 2011. Colored intervals represent event-related activities. The first procession takes place between 11:00 and 12:30 (Interval **(a)**). The second procession takes place between 13:00 and 14:30 (Interval **(b)**). The firework display takes place between 17:00 and 17:30 (Interval **(c)**). About the event: A new Lord Mayor of the City of London is appointed every year and this public parade is organized to celebrate his inauguration. The Lord Mayor participates in a procession from the City of London to the Royal Courts of Justice in the City of Westminster. As in the Middle Ages, he is accompanied by military displays, marching bands, acrobats, dancers, displays of pomp and charity and symbols of London's ancient strength and resolve. The annual one-day event attracts about half a million spectators each year and is one of the City's longest established and best known annual events dating back to 1535. The event starts at 11:00 and the processional route goes from the Mansion House via Bank, St. Paul's Cathedral and Fleet Street to the Aldwych; the tail of the procession will reach the Royal Courts at about 12.30. There is a short break during the ceremony, then the whole procession sets off again at 13:00 to take the new Lord Mayor back to Mansion House. The procession finally ends at about 14:30 when the last floats reach the City.
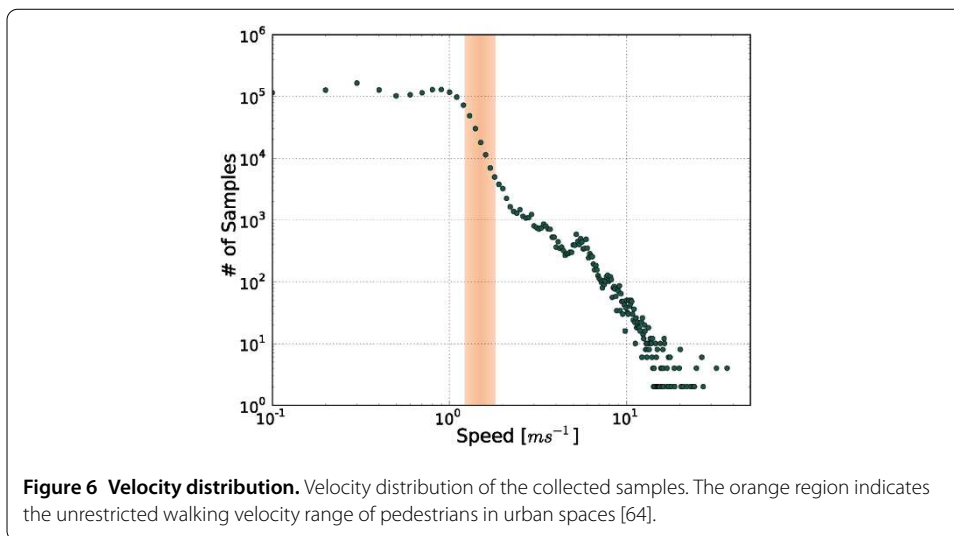
Figure 4 Heat map of user distribution. Map of the data recording zone. Superimposed is a heat map representation of the distribution of all recorded location updates. The 'hotter' an area is colored (*i.e.* the more yellow it is), the more data points have been colected in this area. It is visible that more data has been collected around locations where event-related activities took place.

Figure 5 Spatio-temporal distribution of users. Spatio-temporal distribution of user: **(a)** Distribution of users during the first procession between 11:00 and 12:30 (Interval (a)); **(b)** Distribution during the second procession between 13:00 and 14:30 (Interval (b)); **(c)** Distribution during the firework display between 17:00 and 17:30 (Interval (c)); **(d)** Distribution during the break between 14:30 and 17:00. It is visible that people amass along streets where the processions take place and around the river basin during the fireworks. During the break, however, the accumulation is much lower and concentrated around bus and metro stations.

during the firework display (Interval (c)). Figure 5(d) shows the distribution of reported locations during the break between 14:30 and 17:00. Although temporal information is not present, these heat maps reveal an expected spatial distribution of event attendees: people amass along streets where the processions take place and around the river basin during the fireworks. During the break, however, the accumulation is much lower and concentrations around bus and metro stations are visible.

Figure 6 Velocity distribution. Velocity distribution of the collected samples. The orange region indicates the unrestricted walking velocity range of pedestrians in urban spaces [64].
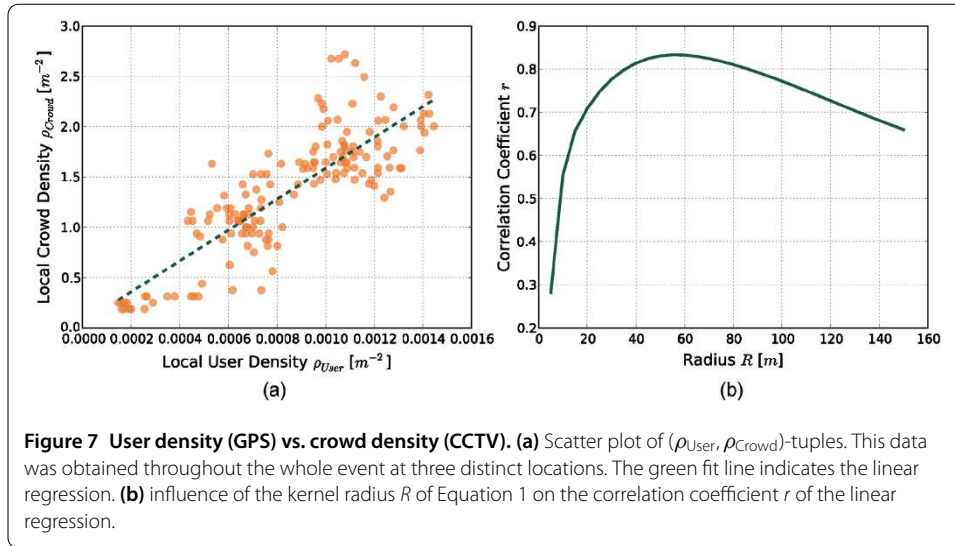
5.2 Velocity distribution

App users do not necessarily walk around by foot but may travel by any means of available transportation. By recording a user's location, the GPS sensor also provides the current velocity the device travels. Figure 6 shows the velocity distribution of the collected data. The orange-colored area indicates the walking velocity range of pedestrians in urban spaces. The mean value is 1.47 ms^{-1} with a variance of 0.3 ms^{-1} according to *Willis et al.* [64]. Walking velocity is affected by cultural influences, demographics and even time of the day and weather conditions. However, these influences lie within the indicated area. The plot reveals that the majority of the collected samples were recorded at a velocities between 0 ms^{-1} and 2 ms^{-1} while only a few data samples were recording at higher velocities. In the following, we are interested in pedestrian dynamics and hence, unless stated otherwise, we only consider data samples where the corresponding velocity lies between 0 ms$^{-1} \leq v \leq 1.47 + 0.3$ ms^{-1} (=1.77 ms^{-1}).

5.3 Relation between user density and crowd density

We assume that the spatial distribution of App users corresponds to the actual spatial distribution of event attendees (Assumption 2). This implies that for a given point in time, the ratio of App users to event attendees is constant for every location. To verify this assumption, we compare the actual crowd density at a specific location to the App user density at the same location. The crowd density is obtained from video footage recorded by CCTV video cameras (see Section 4.2). We use recordings from three different locations and for each of these locations defined an area of approximately 10 m^2 within which the pedestrians are manually counted at certain points in time. Given these counts, the crowd density ρ_{Crowd} is obtained by dividing the number of people N in the area by the size A of the area. Hence:

$$\rho_{\text{Crowd}} = \frac{N}{A}. \tag{4}$$

The corresponding user density ρ_{User} is obtained from the GPS location data using Equation 1. Figure 7(a) shows a scatter plot of the $(\rho_{\text{User}}, \rho_{\text{Crowd}})$-tuples. In total, we obtained 154 density tuples.
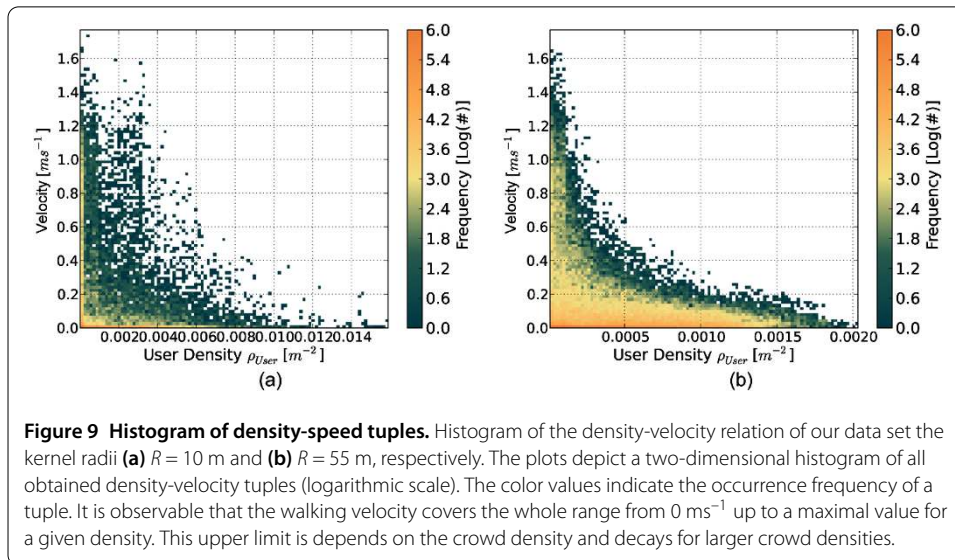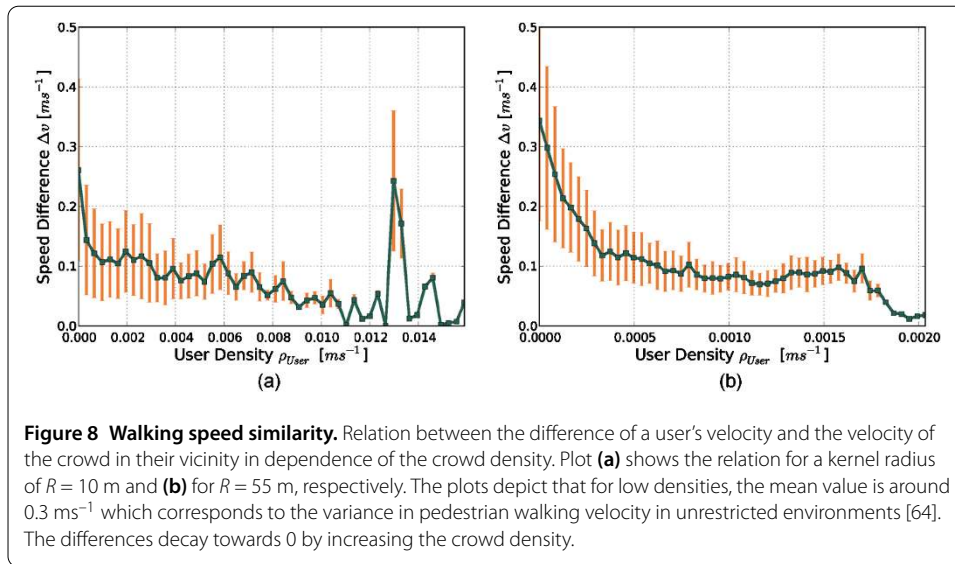
Figure 7 User density (GPS) vs. crowd density (CCTV). (a) Scatter plot of (ρ_{User}, ρ_{Crowd})-tuples. This data was obtained throughout the whole event at three distinct locations. The green fit line indicates the linear regression. **(b)** influence of the kernel radius R of Equation 1 on the correlation coefficient r of the linear regression.

To fulfill Assumption 1, we assume a linear relation between ρ_{Crowd} and ρ_{User}. With a linear regression analysis, we can assess the quality of the linear relation. The linear regression is depicted in Figure 7(a). The user density ρ_{User} depends on the kernel radius R of Equation 1. To understand the influence, we vary the kernel radius R between 5 m < R < 150 m. Figure 7(b) depicts the influence of the kernel radius on the correlation between the crowd density and the user density. We obtain a low correlation coefficient for small values of R. The correlation coefficient increases to a maximum of $r = 0.833$ for $R = 55$ m followed by a decline for larger values of R. The observed behavior can be explained in the following way: This variation is getting smoothed out for larger values of R as the area to determine the density is increased. Hence, small variations in the number of available sample points do not affect the density estimation as greatly resulting in lower variations. By exceeding some value of R, the considered area is so large that the estimated density does not capture the local variation anymore. Local variations are smoothed out and large deviations between the user density and the crowd density can be observed. This causes a drop in the correlation coefficient.

A further error might be introduced by the localization errors due to sub-optimal GPS fixes in urban spaces, where often only a limited number of GPS satellites are visible at the street level. It has been shown in [65] that this error is lower than 24 m for 95% of all samples recorded in urban spaces and that the median error is 8 m.

5.4 Behavioral similarity with respect to density

We assess whether Assumption 3 holds by comparing a user's own velocity to the velocity of their neighbors. For this we determine a user's location and velocity and compare it to the crowd velocity at this location. We calculate the crowd velocity at the user's location using Equation 2 without including the user's own velocity. The velocity difference Δv_k is given by the difference between the user's velocity and the crowd velocity. Hence,

$$\Delta v_k = \left| v_k - \frac{\sum_{i \in \{N \setminus k\}} |\vec{v}_i| \exp[-\|\vec{r}_i(t) - \vec{r}\|^2/R^2]}{\sum_{i \in \{N \setminus k\}} \exp[-\|\vec{r}_i(t) - \vec{r}\|^2/R^2]} \right|, \tag{5}$$

Figure 8 Walking speed similarity. Relation between the difference of a user's velocity and the velocity of the crowd in their vicinity in dependence of the crowd density. Plot **(a)** shows the relation for a kernel radius of $R = 10$ m and **(b)** for $R = 55$ m, respectively. The plots depict that for low densities, the mean value is around 0.3 ms^{-1} which corresponds to the variance in pedestrian walking velocity in unrestricted environments [64]. The differences decay towards 0 by increasing the crowd density.

Figure 9 Histogram of density-speed tuples. Histogram of the density-velocity relation of our data set the kernel radii **(a)** $R = 10$ m and **(b)** $R = 55$ m, respectively. The plots depict a two-dimensional histogram of all obtained density-velocity tuples (logarithmic scale). The color values indicate the occurrence frequency of a tuple. It is observable that the walking velocity covers the whole range from 0 ms^{-1} up to a maximal value for a given density. This upper limit is depends on the crowd density and decays for larger crowd densities.

with ν_k the velocity of user k and N the set of all users. We calculate the velocity difference at each time step for each user together with the local density at that location. The two plots in Figure 8 show the obtained relationship by plotting the velocity difference versus the user density. Plot (a) is obtained with a kernel parameter of $R = 10$ m and (b) with $R = 55$ m, respectively. We see that in both cases, for small densities, the mean value is around 0.3 ms^{-1} which corresponds to the variance in pedestrian walking velocity in unrestricted environments [64]. Additionally, a trend can be observed that the velocity differences tend to get smaller for larger densities. This supports Assumption 3.

5.5 The fundamental diagram: relation between density and velocity

We want to investigate towards which extent the density-velocity relation found in our data set corresponds to existing fundamental diagram models. Figure 9(a) and Figure 9(b) show a histogram of the density-velocity relation for a kernel radius of $R = 10$ m and $R = 55$ m, respectively. To obtain these plots, we divided time into intervals of one second and

calculated for each interval t and for each user that was active in this interval the local density $\rho(\vec{r}, t)$ using Equation 1 and the crowd velocity $v(\vec{r}, t)$ using Equation 2. The plots depict a two-dimensional histogram of all obtained density-velocity tuples (logarithmic scale). The color values indicate the occurrence frequency of a tuple. The two plots reveal some general aspects of the density-velocity relation found in our data sets:

- both plots exhibit a clear trend that with higher densities, the velocity range decreases;
- for low densities, the whole walking velocity range between 0 ms^{-1} and 1.77 ms^{-1} is observed;
- low velocity values can be observed for all densities.

By comparing the obtained results to the density-velocity relation discussed in Section 2.3.2, we see that our data does not look like the plot of the function provided by Weidmann. Our data is scattered across a region as opposed to the bijective mapping of the fundamental diagram. This difference can be explained as follows: The model derived by Weidmann assumes that the pedestrians want to reach a target location. This assumption is not given in our situation. Not every pedestrian has a target location to reach and might decide to walk with his own pace or even decides to stand still. Thus, we can observe walking velocities covering the whole range from 0 ms^{-1} up to a maximal value for a given density. It is, however, observable that this maximal value depends on the crowd density and decreases for higher densities. Therefore, we can conclude that the crowd density value at a given location imposes a restriction on the maximal walking velocity that is possible.

5.6 Calibration of crowd density estimates

Based on the findings deduced in the previous section, we introduce and evaluate a methodology to estimate a crowd density from the spatial distribution of App users. Our method relies on Assumption 2. Section 5.3 shows the existence of a linear relation between the crowd density and the user density. By knowing the parameters of the linear regression, a crowd density can be estimated from the user density. The regression parameters, however, are unknown. Thus, a calibration method is required to obtain these parameters.

5.6.1 Calibrating the spatial distribution of App users to obtain crowd density estimates

By using Equation 1, we obtain a local user density ρ_{User} from the spatial distribution of App users. Making use of the linear relation, we obtain a local crowd density estimation $\hat{\rho}_{\text{Crowd}}$ from the measured local user density ρ_{User}:

$$\hat{\rho}_{\text{Crowd}}(\vec{r}, t) = \frac{m}{k} \rho_{\text{User}}(\vec{r}, t) + \frac{q}{k}, \tag{6}$$

where m, q and k are unknown regression parameters and depend on the ratio of App users to event attendees.

Section 2.3.2 presents Weidmann's analytical equation to model the fundamental diagram (Equation 3). This equation describes the crowd speed as a function of the crowd density. It can be transformed so that the crowd density is a function of the crowd speed:

$$\rho_{\text{Crowd}}(\vec{r}, t, v) = \frac{\gamma \cdot \rho_{\max}}{\rho_{\max} \cdot \ln\left(\frac{-v_0}{v(\vec{r},t)-v_0}\right) + \gamma} \quad \text{with } v(\vec{r}, t) \neq v_0. \tag{7}$$

The speed of the crowd $v(\vec{r}, t)$ is obtained using Equation 2. Hence, we can obtain a crowd density estimates $\tilde{\rho}_{\text{Crowd}}$ by combining Equation 2 and Equation 7. The parameters $\rho_{\max}$ and v_0 are cultural dependent and can be taken from literature (*e.g.* [48, 53, 55]). The fitting parameter γ, however, remains unknown.

For a given time at a given location, Equation 6 and Equation 7 should provide the same crowd density estimates $\hat{\rho}_{\text{Crowd}}$ and $\tilde{\rho}_{\text{Crowd}}$. Hereby, Equation 6 considers the local user density and Equation 7 the local crowd speed. We define an error measure e:

$$e = \left(\hat{\rho}_{\text{User}}(\vec{r}, t) - \rho_{\text{User}}(\vec{r}, t) \right)^2 \tag{8}$$

with

$$\hat{\rho}_{\text{User}}(\vec{r}, t) \overset{(6)}{=} \frac{k \cdot \hat{\rho}_{\text{Crowd}}(v) - q}{m} = \frac{k \cdot \tilde{\rho}_{\text{Crowd}}(v) - q}{m}$$

$$\overset{(7)}{=} \frac{\frac{k \cdot \gamma \cdot \rho_{\max}}{\rho_{\max} \cdot \ln(\frac{-v_0}{v(\vec{r},t) - v_0}) + \gamma} - q}{m}, \quad m \neq 0. \tag{9}$$

The missing calibration parameters m, q and γ can now be found by minimizing the error e with a least square method. The minimization criteria we used is
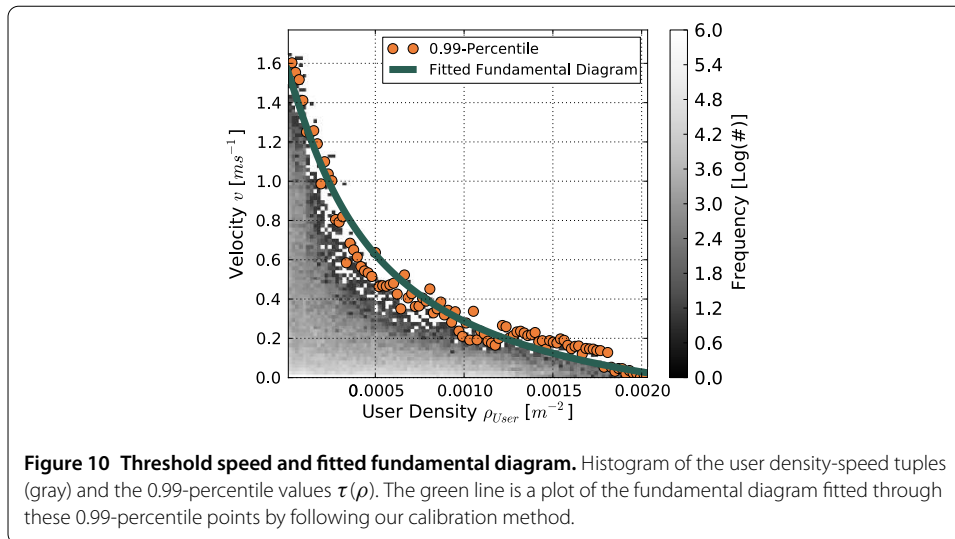
$$S(m, q, \gamma) = \sum_{i=1}^{N} \left[\hat{\rho}_{\text{User}}(m, q, \gamma, v, k) - \rho_{\text{User}} \right]^2. \tag{10}$$

5.6.2 Modeling the fundamental diagram from the recorded density-speed information

With the previous approach, we can obtain the optimal calibration parameters m and q by using Weidmann's equation to fit the user density to the corresponding crowd speed. However, the density-speed tuples do not represent the fundamental diagram well as there is a great amount of variation in the walking behavior of pedestrians (Section 5.5). We found in our data set that pedestrians walk with a speed between 0 ms^{-1} and a density-dependent upper limit. We consider this upper limit as the speed with which pedestrians' walking behavior gets restricted by the surrounding crowd. Increasing the personal walking speed would conflict with the social forces acting on a pedestrian [27]. Our assumption is that pedestrians walking with the upper limit speed for a given density behave according to the fundamental diagram. Hence, we perform a calibration with only these upper limit values. To obtain the upper limit values, we introduce $\tau(\rho)$, the 0.99-percentile value. $\tau(\rho)$ is the threshold speed for a given density ρ for which 99% of all measured speed values are smaller. Figure 10 shows again the frequency plot of the $(\rho_{\text{User}}(\vec{r}, t), v(\vec{r}, t))$-tuples together with the 0.99-percentile values $\tau(\rho)$. These percentile values $\tau(\rho)$ can now be used to minimize Equation 10 to obtain the calibration parameters m and q. The green curve in Figure 10 shows the calibrated fundamental diagram. Hereby, we set $\rho_{\max} = 3.5$ m^{-2} (According to Weidmann [48]) and $v_0 = 1.77$ ms^{-1} (according to *Willis et al.* for UK [64]). Table 3 lists the calibration parameters obtained by our minimization process for different kernel radii R.

5.6.3 Evaluation of the calibration methodology

To gain insight into the accuracy of our calibration methodology, we calibrate all user density measure ρ_{User} where a CCTV-based reference crowd density is available. This is the

Figure 10 Threshold speed and fitted fundamental diagram. Histogram of the user density-speed tuples (gray) and the 0.99-percentile values $\tau(\rho)$. The green line is a plot of the fundamental diagram fitted through these 0.99-percentile points by following our calibration method.

Table 3 Overview of calibration parameter

R	m	q	γ	k
10	550	0	1.6	1.7
15	624	0.1	1.1	1.4
20	757	0	0.6	1.5
25	1,176	0.1	0.6	1.9
30	1,262	0.2	0.5	1.7
35	618	0.1	0.5	0.7
40	1,972	0.3	0.5	1.9
45	1,312	0.2	0.5	1.1
50	1,339	0.2	0.5	1.0
55	1,177	0.2	0.5	0.8

Calibration parameters obtained through our calibration method for different kernel radii.

same data as used in Section 5.3. We compare the outcome to the CCTV-based reference data. Ideally, the estimated crowd density $\hat{\rho}_{\text{Crowd}}$ obtained from the calibrated App user distribution should be identical to the observed crowd density ρ_{Crowd} from the video footage. We apply a linear regression trough the data tuples to understand the calibration accuracy. Figure 11 shows the linear regressions for different kernel radii. A perfect regression would correspond to the diagonal axis. We see that all regressions are situated around the diagonal axis.

We perform a residual analysis to assess the appropriateness of the chosen model. A residual is defined as follows:

$$\epsilon = \hat{\rho}_{\text{Crowd}} - \rho_{\text{Crowd}}. \tag{11}$$

Figure 12(a) is a plot of the residuals for the kernel radii $R = 10$ m and $R = 55$ m dependent on the crowd density. Figure 12(b) shows the normal probability plot. The normal probability plot helps to determine whether or not it is reasonable to assume that the random errors in a statistical process can be assumed to be drawn from a normal distribution. The normal probability plot shows a strongly linear pattern. With a linear regression fitted through the data (dashed lines), we obtain a correlation coefficient of 0.985 for $R = 10$ m and 0.969 for $R = 55$ m, respectively. These correlation coefficients indicate that there are

Figure 11 Calibration regression. Linear regressions of the calibrated crowd densities $\hat{\rho}_{Crowd}$ and the crowd densities ρ_{Crowd} obtained from video footage. Results are shown for different kernel radii R. A perfect regression would correspond to the diagonal axis. Additionally, the figure also shows the scatter plot of the calibrated data points for the case of $R = 55$ m.

Figure 12 Residual analysis. (a) Residual plot of the estimated crowd density to the actual crowd density for the two kernel radii $R = 10$ m and $R = 55$ m. **(b)** The normal probability plot of the residuals. The good linearity of the regression supports the chosen model. **(c)** The histogram of the residuals shows a normal distribution.

Table 4 Correlation coefficient and calibration error

R [m]	r	σ (RMSE) [m^{-2}]
10	0.56	0.54
15	0.66	0.46
20	0.71	0.51
25	0.75	0.46
30	0.78	0.41
35	0.80	0.37
40	0.81	0.35
45	0.82	0.34
50	0.83	0.35
55	0.83	0.36

Results of the crowd density calibration method by comparing the obtained crowd density estimates to ground truth information extracted from CCTV video footage. The density estimates rely on the kernel radius R which is used to determine the user density. Results for different values of R are listed. The correlation coefficient r indicates how well the crowd density estimates correlate to the ground truth measures. The crowd density estimation error σ is the average estimation error.

only minor deviations from the line fit to the points on the probability plot. Hence, the chosen model appears to be suitable to model the data. This finding is also supported by the histogram depicted in Figure 12(c) which shows that the residuals have a normal distribution.

To understand how well we can estimate the crowd density from the distribution of App users, we determine the overall calibration error by calculating the root mean squared error (RMSE) σ as follows:

$$\sigma = \sqrt{\frac{\sum_i^N (\rho_{\text{Crowd}} - \hat{\rho}_{\text{Crowd}})^2}{N}}$$

$$= \sqrt{\frac{\sum_i^N (\rho_{\text{Crowd}} - (\frac{m}{k} \cdot \rho_{\text{User}} + \frac{q}{k}))^2}{N}}. \tag{12}$$

Table 4 lists σ for different kernel radii. The table also lists the obtained correlation coefficients r of a linear regression through the actual crowd density ρ_{Crowd} and the estimation $\hat{\rho}_{\text{Crowd}}$.

Given all these findings, we conclude:

- The residual analysis reveals that the error is normal distributed which suggests that the chosen model fits the data well and that the error is not introduced by the model but inherently present in the data,
- we achieve a correlation coefficient of $r = 0.56$ for $R = 10$ m and $r = 0.83$ for $R = 55$ m, respectively. This implies that there is some predicting power for obtaining a crowd density estimation, and
- the calibration error is $\sigma = 0.54$ m^{-2} for $R = 10$ m and $\sigma = 0.36$ m^{-2} for $R = 55$ m, respectively.

6 Conclusion

A participatory sensing approach for crowd monitoring faces a major limitation: Participation is based on a voluntary base. Regardless of the incentivization strategy, we expect that only a small fraction of all attendees of a mass gathering is being tracked. This makes it challenging to conclude about the crowd density. This work addressed this limitation. We presented a methodology which allows to infer a crowd density even if only a small

number of crowd members is being tracked. The principle behind our methodology is that the walking speed of pedestrians depends on the crowd density. By measuring the location and speed, we can calibrate the distribution of tracked pedestrians to the distribution of all attendees of a mass gathering using the fundamental diagram. With this, we can infer crowd density estimates.

We used a data set recorded during a city-scale mass gathering to evaluate our methodology. We compared crowd density estimates to ground truth information obtained from video footage: For a kernel radius of $R = 55$ m, the average calibration error is 0.36 m^{-2}. Further, a correlation coefficient of 0.83 indicates that a linear relation between the crowd density and the user density can be assumed. The residual analysis revealed that the model fits the data well.

Besides these results, the work presents another finding: We could verify that the walking speed of pedestrians depends on the crowd density. Hereby, we found a similar relation between the speed of a crowd and the density as related work suggests. In particular, we identified a crowd density dependent upper limit speed with which pedestrians move through urban spaces. These upper speed limit values follow existing fundamental diagram models closely.

There are several factors to consider:

- The reason for not reaching a higher correlation coefficient than the maximum value of $r = 0.83$ might stem from the unequal spatial distribution of App users and event attendees at certain time steps. However, there are also other factors: It was sometimes difficult to count the correct number of attendees in the predefined area from the video footage as some pedestrians were occluded by others. Therefore, the crowd density extracted from the video is also error-prone.
- We obtained the highest correlation coefficient and lowest calibration error for a kernel radius $R = 55$ m. This is a large radius to infer local characteristics. We believe this is due to the sparsity in our data set. We were tracking less than 1% of all attendees. A smaller kernel radius could provide more accurate local crowd information [20] but would require a much larger user base. Providing more attractive incentives, making the App available on different mobile platforms and having a good advertisement campaign in place could stimulate a higher participation.
- We obtained best results with a radius of 55 m. This seems to be like a big area to cover for monitoring crowd. However, as we use a Gaussian weighting scheme to calculate our measures, the influence of the users decays rapidly the further away they are from the center of the circle. Further, we believe that this radius can be smaller by having a larger ratio of App users.

The location sampling rate of 1 Hz was chosen to capture as much of the pedestrian dynamics as possible. However, such a high sampling rate is very energy consuming. Besides privacy considerations, also the heavy battery consumption of such an App might have a detrimental effect on participation. Therefore, it is important to incorporate an efficient energy conserving sampling strategy. This can be achieved by lowering the sampling frequency but also by only reading location updates from GPS if needed. Hereby, low-power acceleration sensors can help to determine if a user is stationary or not and only switch on the GPS if motion is being detected.

Another important issue that has not been addressed in this work is to obtain a confidence measure giving indication about the reliability of the inferred crowd density. It may

be that due to a small percentage of users compared to the total number of attendees, the inferred crowd density may even become null. Hereby, a plausibility check *e.g.* by comparing the active number of users to a roughly estimated number of attendees by the security personnel could give confidence about the inferred crowd density.

This work is one of the first addressing the challenges arising by crowd sensing through a participatory sensing approach with smartphones. We believe the results are promising to stimulate successive contributions. In particular, we see the following next steps to investigate some of the aspects not addressed in this work:

- We evaluated our approach on data from only one mass gathering. To generalize the findings, our method has to be applied to data collected during different mass gatherings and the results have to be compared. The type of the gathering and cultural aspects may have an influence.
- A sensitivity analysis investigating the relation between the ratio of App users and the accuracy of crowd density estimation helps to understand how many pedestrian need to be tracked to obtain a significant estimation accuracy.
- An evaluation of the online performance of our method reveals the required amount of data to estimate a crowd density. The required amount of data is closely connected to the required amount of pedestrians. These two aspects should be investigated jointly.
- We used the analytical model of Weidmann to represent the fundamental diagram. As noted in Section 2.3.2, other models exist which consider additional information. The suitability of alternative models for our calibration method remains to be investigated.
- A possible demographic bias in our App usage was not taken into consideration. However, such factors influence the behavior of pedestrians. Considering the age or gender distribution or the cultural background could further tune the model parameters.
- We did not consider to include spatial characteristics into our model. As the behavior of pedestrians depends on the architectural configuration, such information could be considered to increase the estimation accuracy.

This work shows on the example of crowd density that a participatory sensing approach can give insight into crowd characteristics and provide information relevant to assess the criticality of a situation during city-scale mass gatherings. Given our results and the many advantages of on-device localization (localization accuracy, user control over privacy, multitude of sensor modalities, low deployment cost, *etc.*), we suggest that smartphones are a viable tool for crowd monitoring.

Author details
[1]Wearable Computing Laboratory, ETH Zürich, Zürich, Switzerland. [2]Embedded Intelligence Group, DFKI Kaiserslautern, Kaiserslautern, Germany. [3]Complexity Research Programme, London School of Economics and Political Science, London, United Kingdom.

References
1. BBC News (2011) Royal wedding: in numbers. http://www.bbc.co.uk/news/uk-13248642
2. Clarin.com (2010) Dos millones de personas, en un cierre inolvidable. http://edant.clarin.com/diario/2010/05/25/um/m-02199900.htm
3. RTVE (2010) Dos millones de personas recibieron a la Selección Española de Fútbol en Madrid. http://www.rtve.es/alacarta/videos/television/dos-millones-personas-recibieron-a-seleccion-espanola-futbol-madrid/828627/
4. Helbing D, Mukerji P (2012) Crowd disasters as systemic failures: analysis of the love parade disaster EPJ Data Sci 1:7. http://www.epjdatascience.com/content/1/1/7
5. Sime J (1995) Crowd psychology and engineering. Saf Sci 21:1-14
6. Tarlow P (2002) Event risk management and safety, vol 4. Wiley, New York
7. Getz D (2007) Event studies: theory, research and policy for planned events. Elsevier, Amsterdam
8. Johansson A, Batty M, Hayashi K, Al Bar O, Marcozzi D, Memish ZA (2012) Crowd and environmental management during mass gatherings. Lancet Infect Dis 12(2):150-156
9. Helbing D, Buzna L, Johansson A, Werner T (2005) Self-organized pedestrian crowd dynamics: experiments, simulations, and design solutions. Transp Sci 39:1-24
10. Hughes RL (2003) The flow of human crowds. Annu Rev Fluid Mech 35:169-182. http://www.annualreviews.org/doi/abs/10.1146/annurev.fluid.35.101101.161136.
11. Smelser NJ (1963) Theory of collective behavior
12. Nicholson C, Roebuck B (1995) The investigation of the Hillsborough disaster by the health and safety executive. Saf Sci 18(4):249-259
13. Au S, Great Britain H, Staff SE, Health GB, Executive S, Ltd RC (1993) Managing crowd safety in public venues: a study to generate guidance for venue owners and enforcing authority inspectors. HSE contract research report, HSE Books. http://books.google.ch/books?id=3osbPwAACAAJ
14. Wirz M, Franke T, Roggen D, Mitleton-Kelly E, Lukowicz P, Tröster G (2012) Inferring and visualizing crowd conditions by collecting GPS location traces from pedestrians' mobile phones for real-time crowd monitoring during city-scale mass gatherings. In: Collaboration technologies and infrastructures (WETICE), 21st international conference on. IEEE Press, New York
15. Batty M, Desyllas J, Duxbury E (2003) The discrete dynamics of small-scale spatial events: agent-based models of mobility in carnivals and street parades. Int J Geogr Inf Sci 17(7):673-697
16. Lee R, Hughes R (2005) Exploring trampling and crushing in a crowd. J Transp Eng 131:575
17. Hsieh Y, Ngai K, Burkle F Jr, Hsu E (2009) Epidemiological characteristics of human stampedes. Disaster Med Public Health Prep 3(4):217
18. Burkle F, Hsu E (2011) Ram Janki temple: understanding human stampedes. Lancet 377(9760):106-107
19. Krausz B, Bauckhage C (2012) Loveparade 2010: automatic video analysis of a crowd disaster. Comput Vis Image Underst 116(3):307-319
20. Helbing D, Johansson A, Al-Abideen H (2007) Dynamics of crowd disasters: an empirical study. Phys Rev E 75(4):046109
21. Krausz B, Bauckhage C (2011) Analyzing pedestrian behavior in crowds for automatic detection of congestions. In: Computer vision workshops (ICCV workshops), 2011 IEEE international conference on. IEEE Press, New York, pp 144-149
22. Johansson A, Helbing D, Al-Abideen HZ, Al-Bosta S (2008) From crowd dynamics to crowd safety: a video-based analysis. ArXiv e-prints
23. Fruin J (1981) Crowd disasters - a systems evaluation of causes and countermeasures. Inc. US National Bureau of Standards, pub. NBSIR, 81-3261
24. Gong S, Loy CC, Xiang T (2011) Security and surveillance. In: Moeslund TB, Hilton A, Krüger V, Sigal L (eds) Visual analysis of humans. Springer, London
25. Jacques J Jr, Musse S, Jung C (2010) Crowd analysis using computer vision techniques. IEEE Signal Process Mag 27:66-77
26. Mehran R, Oyama A, Shah M (2009) Abnormal crowd behavior detection using social force model. In: Computer vision and pattern recognition. IEEE Press, New York
27. Helbing D, Molnar P (1995) Social force model for pedestrian dynamics. Phys Rev E 51:4282-4286
28. Steffen B, Seyfried A (2010) Methods for measuring pedestrian density, flow, speed and direction with minimal scatter. Phys A, Stat Mech Appl 389(9):1902-1910
29. Eagle N, Pentland A (2006) Reality mining: sensing complex social systems. Pers Ubiquitous Comput 10(4):255-268
30. LaMarca A, Chawathe Y et al (2005) Place lab: device positioning using radio beacons in the wild. In: Pervasive computing
31. Calabrese F, Colonna M, Lovisolo P, Parata D, Ratti C (2011) Real-time urban monitoring using cell phones: a case study in Rome. IEEE Trans Intell Transp Syst 12:141-151
32. Becker R, Caceres R, Hanson K, Loh J, Urbanek S, Varshavsky A, Volinsky C (2011) A tale of one city: using cellular network data for urban planning. IEEE Pervasive Comput 10(4):18-26
33. Couronne T, Olteanu Raimond A, Smoreda Z (2011) Looking at spatiotemporal city dynamics through mobile phone lenses. In: Network of the future, international conference on the. IEEE Press, New York
34. Calabrese F, Pereira F, Di Lorenzo G, Liu L, Ratti C (2010) The geography of taste: analyzing cell-phone mobility and social events. In: Pervasive computing. Lecture notes in computer science, vol 6030. Springer, Berlin, pp 22-37
35. Kim D, Kim Y, Estrin D, Srivastava M (2010) Sensloc: sensing everyday places and paths using less energy. In: Proc. of the 8th ACM conference on embedded networked sensor systems. ACM, New York

36. Van Diggelen F (2009) A-GPS: assisted GPS, GNSS, and SBAS. Artech House, Norwood
37. Azizyan M, Constandache I, Choudhury RR (2009) SurroundSense: mobile phone localization via ambience fingerprinting. In: Proceedings of the 15th annual international conference on mobile computing and networking, MobiCom '09. ACM, New York, pp 261-272
38. Koshak N, Fouda A (2008) Analyzing pedestrian movement in mataf using gps and gis to support space redesign. In: The 9th international conference on design and decision support systems in architecture and urban planning
39. Versichele M, Neutens T, Delafontaine M, de Weghe NV (2012) The use of bluetooth for analysing spatiotemporal dynamics of human movement at mass events: a case study of the Ghent festivities. Appl Geogr 32(2):208-220
40. Bandini S, Federici ML, Manzoni S (2007) A qualitative evaluation of technologies and techniques for data collection on pedestrians and crowded situations. In: Proceedings of the 2007 summer computer simulation conference, SCSC. Society for Computer Simulation International, San Diego, pp 1057-1064
41. Marana A, Da Fontoura Costa L, Lotufo R, Velastin S (1999) Estimating crowd density with Minkowski fractal dimension. In: Acoustics, speech, and signal processing, IEEE international conference on, vol 6, pp 3521-3524
42. Ma R, Li L, Huang W, Tian Q (2004) On pixel count based crowd density estimation for visual surveillance. In: Cybernetics and intelligent systems, IEEE conference on, vol 1. IEEE Press, New York, pp 170-173
43. Wu X, Liang G, Lee K, Xu Y (2006) Crowd density estimation using texture analysis and learning. In: Robotics and biomimetics, ROBIO'06, IEEE international conference on. IEEE Press, New York, pp 214-219
44. Jones M, Snow D (2008) Pedestrian detection using boosted features over many frames. In: Pattern recognition, ICPR 2008, 19th international conference on. IEEE Press, New York, pp 1-4
45. Brostow G, Cipolla R (2006) Unsupervised Bayesian detection of independent motion in crowds. In: Computer vision and pattern recognition, IEEE computer society conference on, vol 1. IEEE Press, New York, pp 594-601
46. Reades J, Calabrese F, Sevtsuk A, Ratti C (2007) Cellular census: explorations in urban data collection. IEEE Pervasive Comput 6(3):30-38
47. Morrison A, Bell M, Chalmers M (2009) Visualisation of spectator activity at stadium events. In: Information visualisation, 13th international conference. IEEE Press, New York, pp 219-226
48. Weidmann U (1992) Transporttechnik der Fussgänger: Transporttechnische Eigenschaften des Fussgängerverkehrs (Literaturauswertung). IVT, Zürich
49. Helbing D (2009) Derivation of a fundamental diagram for urban traffic flow. Eur Phys J B, Condens Matter Complex Syst 70(2):229-241
50. Jelić A, Appert-Rolland C, Lemercier S, Pettré J (2011) Properties of pedestrians walking in line-fundamental diagrams. Arxiv preprint arXiv:1111.5708
51. Daamen W, Hoogendoorn S (2003) Experimental research of pedestrian walking behavior. Transp Res Rec 1828:20-30
52. Zhang J, Klingsch W, Schadschneider A, Seyfried A (2012) Ordering in bidirectional pedestrian flows and its influence on the fundamental diagram. J Stat Mech Theory Exp 2012:P02002
53. Chattaraj U, Seyfried A, Chakroborty P (2009) Comparison of pedestrian fundamental diagram across cultures. ArXiv e-prints
54. Smith R (1995) Density, velocity and flow relationships for closely packed crowds. Saf Sci 18(4):321-327
55. Johansson A (2009) Constant-net-time headway as a key mechanism behind pedestrian flow dynamics. Phys Rev E 80(2):026120
56. Seyfried A, Steffen B, Lippert T (2006) Basics of modelling the pedestrian flow. Phys A, Stat Mech Appl 368:232-238
57. Fang Z, Lo S, Lu J (2003) On the relationship between crowd density and movement velocity. Fire Saf J 38(3):271-283
58. Seyfried A, Steffen B, Klingsch W, Boltes M (2005) The fundamental diagram of pedestrian movement revisited. J Stat Mech Theory Exp 2005:P10002
59. Schadschneider A, Klingsch W, Klüpfel H, Kretz T, Rogsch C, Seyfried A (2008) Evacuation dynamics: empirical results, modeling and applications. Arxiv preprint arXiv:0802.1620
60. Wiseman R (2008) Quirkology: the curious science of everyday lives. Pan, London
61. Wirz M, Roggen D, Tröster G (2010) User acceptance study of a mobile system for assistance during emergency situations at large-scale events. In: Human-centric computing (HumanCom), 3rd international conference on. IEEE Press, New York, pp 1-6
62. Warburton K, Lazarus J (1991) Tendency-distance models of social cohesion in animal groups. J Theor Biol 150(4):473-488
63. Moussaid M, Garnier S, Theraulaz G, Helbing D (2009) Collective information processing and pattern formation in swarms, flocks, and crowds. Top Cogn Sci 1(3):469-497
64. Willis A, Gjersoe N, Havard C, Kerridge J, Kukla R (2004) Human movement behaviour in urban spaces: implications for the design and modelling of effective pedestrian environments. Environ Plan B, Plan Des 31(6):805-828
65. Wirz M, Schläpfer P, Kjærgaard M, Roggen D, Feese S, Tröster G (2011) Towards an online detection of pedestrian flocks in urban canyons by smoothed spatio-temporal clustering of GPS trajectories. In: Proceedings of the 3rd ACM SIGSPATIAL international workshop on location-based social networks. ACM, New York