Performance Evaluation of Mobile Crowdsensing for Event Detection

Matthias Hirth, Stanislav Lange, Michael Seufert, Phuoc Tran-Gia

University of Würzburg, Institute of Computer Science, Chair of Communication Networks, Würzburg, Germany
{matthias.hirth | stanislav.lange | seufert | trangia}@informatik.uni-wuerzburg.de

Abstract—**Crowdsensing offers a cost effective way to collect large amounts of sensor data. However, in contrast to fixed sensor deployments, the spatial distribution of the sensors can hardly be influence, as the sensors are carried by participants of the crowdsensing system. This in turn raises the question about the performance of such systems with respect to the detection probability and detection time of spatial events. In order to address this question, we analyze the performance of such a crowdsensing system by means of simulation. We use the traffic infrastructure of a small size city in Germany and simulate the inhabitants' movement patterns with the well established SUMO mobility generator. Our results show that even if only a small share of inhabitants participates in crowdsensing, events, which have locations that are correlated with the population density, can be easily and quickly detected using such a system. On the contrary, events whose locations are uniformly randomly distributed are much harder to detect using a crowdsensing approach.**

I. Introduction

People have been collecting sensor measurements in urban areas for decades to derive environmental models or to adapt their behavior to changing situations, like traffic routing with respect to the current traffic volumes. In the past, this process of data collection, data analysis, and deduction of models or action guidelines was time consuming and the overall coverage of the sensing information was rather limited due to the required number of dedicated sensing equipment. However, the rise of novel concepts like smart cities creates an increasing demand for fine grained and up-to-date environmental information, which cannot be fulfilled with traditional approaches that solely build on a small number of highly specialized (offline) sensing equipment.

One possibility to tackle this challenge is the usage of a large number of cheap Internet of Things (IoT)-based sensing nodes. Recently many vendors started to offer cheap hardware boards that combine Internet connectivity, low power consumption, and simple programmability. These boards can again be used as basis for customized sensing nodes that continuously deliver real-time environmental data. Another option to collect large amounts of environmental data is crowdsensing. With the still increasing number of smartphones, smart devices, and wearables, a lot of people carry a diverse set of sensory equipment, including, for example, microphones, cameras, brightness sensors, and gyroscopes. Due to the connectivity features of the devices, the sensor information can be made available in almost

real time and can often be further combined with location information, e.g., based on the devices' GPS receiver. Crowdsensing tries to leverage this source of sensing data by directly involving people in the collection of environmental data. Especially due to the low investment costs, as no additional sensor hardware needs to be deployed, crowdsensing is a promising source for sensor information in smart cities.

However, one major drawback of crowdsensing is the missing control of the spatial distribution of the sensor nodes. The sensor nodes, i.e., the smart devices, are carried by the participating citizens and the density of the sensory network is consequently highly correlated with the population density. Considering the daily routines of the crowdsensing participants, e.g., going to work in the morning and returning home at night, both the population density as well as the geographical density of the sensor network even change during the day. With respect to this limitation, the question arises how good the actual sensor coverage of a crowdsensing approach is, and if crowdsensing can be used to reliably detect spatial events, respectively.

In this work, we address this question by using a simulative evaluation of a real-life scenario, in which inhabitants of a small city contribute to a crowdsensing system in order to detect different types of spatial events, which can be correlated or uncorrelated with the density of people. For all event types, we analyze the detection probability, i.e., the share of events that is detected by the crowdsensing users, and also the detection time, i.e., the time between the appearance of the event and its detection. The traffic infrastructure in our example is based on OpenStreet data for the city of Würzburg, Germany and the movement patterns of the crowdsensing participants are generated using the Simulation of Urban Mobility (SUMO) [1]. Our results show that even if only 1% of the 125 000 inhabitants of the city contributes to the crowdsensing system, correlated events can be detected with a very high probability and within a short time after their occurrence. In contrast, uncorrelated events are harder to detect using a crowdsensing based approach and only about 30% of them can be detected in a reasonable about of time.

The remainder of the paper is structured as follows. Section II reviews related work and puts our work in context. Section III details on the methodology, including the generation of the events, the mobility model used for the citizens, and the event detection. Section IV presents and discusses the outcome of our simulation experiments with respect to the detection time and

detection probability. Finally, Section V concludes our paper and points our directions for future research.

II. RELATED WORK

In this section, we first cover general concepts that are relevant in the context of crowdsensing. These include the architecture of such systems, incentives for user participation, goal functions of platform providers, and use cases. Afterwards, we discuss related work that deals with mobile crowdsensing (MCS) in particular, i.e., location specific tasks, spatial coverage, and user mobility. Furthermore, we provide an overview of hybrid systems that merge sensed data from mobile users and fixed sensors.

A. Crowdsensing Systems

The widespread availability of smartphones that are equipped with different sensors as well as cameras paves the way for large scale crowdsensing This enables use cases such as temperature and traffic monitoring [2], WiFi localization [3] as well as characterization [4].

In order to cope with the amount and frequency of information exchange within crowdsensing systems in an efficient manner, several works [5], [6] propose architectural frameworks that standardize the common steps of sensing, transmission, aggregation, processing, and forwarding to applications. Additionally, protocols [7] and applications [3] that address the overhead in terms of energy consumption in the context of MCS lower the barrier for end-user participation.

Crowdsensing service providers strive to optimize two main goals, namely minimizing the cost for sensing and maintaining high quality, reliable data [2]. Furthermore, a high user participation is required for keeping sensed information up-to-date. Therefore, many research initiatives also deal with the topic of incentives [8]. In most cases, these incentives are monetary and are tuned in order to favor honest reports while minimizing expenditures for the provider. Such techniques include reverse auction approaches [9] as well as reputation systems that quantify users' trustworthiness [10]. In this work, however, we assume that users provide honest reports.

B. Mobility and Location Awareness

Similarly to this work, where events occur at specific locations and whose detection requires the presence of nearby users, several works deal with the challenges of location-specific tasks. For example, the authors of [11] address the location uncertainty that arises from users who hide their location due to privacy concerns.

While in our work, users do not stray from their regular path to detect events, the authors of [12] propose the notion of a time budget which can be spent on detours for crowdsensing tasks. Both of the aforementioned works note that the corresponding optimization problems for achieving minimal costs are NP hard and propose heuristic and approximation algorithms to cope with large scale problem instances.

In [13], hybrid sensor deployments are discussed. In this context, crowdsensed data is combined with readings from fixed sensors and cameras to increase the performance in terms of precision and coverage. To reap the benefits of such hybrid systems as well as systems that feature only fixed sensors, the spatial placement of the fixed sensors needs to be optimized [14], [15] and a suitable notion of coverage should be chosen [16]. In our particular case, less frequented areas of the city would be candidates for sensor locations that enable quick event detection despite a low population density.

Finally, not all possible participants might be required to meet the constraints of a particular crowdsensing service. Therefore, the provider might recruit only a subset of users to minimize the costs. Simulation studies with algorithms that solve the corresponding optimization problem [17] as well as case studies [18] demonstrate that such user selection strategies can significantly reduce the payments while maintaining a high service quality. Hence, these techniques could also be used in our event detection context during busy hours when the population density in urban areas allows almost instantaneous detection. Furthermore, social interaction between humans can be used to reduce network overhead by exchanging and aggregating data locally before sending it to the service provider [19].

III. METHODOLOGY

The performance evaluation of mobile crowdsensing focuses on the event detection scenario, i.e., events appear at random times and random locations on a map and have to be detected by the sensors. Ignoring the shielding caused by obstacles, a regular placement of fixed sensors can cover the whole map, which leads to an immediate detection of all events. However, depending on the sensors' detection range, a large number of fixed sensors is needed, which causes high capital and operational expenditures. In this work, the potential savings and trade-offs of mobile crowdsensing are investigated. It is assumed that users do opportunistic crowdsensing, i.e., they move on the map and passively sense their environment. They are not instructed to move to or sense a particular area and might even be completely unaware of the sensing, e.g., if sensing is a background process on their smartphone. They are considered as mobile sensors and, similar to fixed sensors, have a given detection range. Their coverage is determined by their activity and mobility, such that there is a probabilistic availability of sensor measurements in terms of time and location.

The performance evaluation is conducted by means of a discrete event simulation. In the following, the details of the simulation are presented.

A. Event Appearance

In this performance evaluation, $800\,000$ events are simulated and the time of appearance of each event is independently and uniformly distributed over one day. The event location is determined randomly according to one of three methods:

First, *uncorrelated events* are considered whose location of appearance is uncorrelated to the density of people, e.g., rain or lightning strikes. In this case, the event location is uniformly distributed over the map. Second, events might also be correlated to the density of people, e.g., accidents or traffic

jams. Therefore, the location of *correlated events* is distributed identically to the density of people. Finally, *partially correlated events* are events, which depend on people only to some extent, e.g., fire. They are modeled by introducing a percentage p, such that $p\%$ of the events are correlated events, i.e., the location is distributed according to population density, and $(100-p)\%$ are uncorrelated events, i.e., the location is uniformly distributed.

B. User Mobility

To generate movements of users, the Simulation of Urban Mobility (SUMO)[1] [1] is used. SUMO is a free and open traffic simulation suite implemented by Deutsches Zentrum für Luft- und Raumfahrt (DLR, German Aerospace Center). It allows the modeling of intermodal traffic systems including road vehicles, public transport, and pedestrians. Thereby, the highly customizable tool can create mobility traces for arbitrary cities (e.g., based on OpenStreetMap data) and purposes (e.g., traffic light control, emission calculation).

For this study, a pedestrian mobility trace was generated with SUMO for the small city of Würzburg, Germany with a population of $125\,000$. A map of Würzburg of size 8.75 km x 6.05 km was obtained from OpenStreetMap and imported into SUMO as a road network. The mobility simulation spawns a new pedestrian every second, who undertakes a single trip on the map, and then vanishes. For each pedestrian, two edges of the SUMO network, i.e., roads, are selected uniformly random as the start and end of the trip with a maximum distance of 2 km. SUMO performs a fastest-path routing to determine the intermediate edges. Based on the trip definition and a pedestrian model with default parameters[2] (e.g., maximum speed 5.4 km/h), the position of the pedestrian is computed at every second and added to the mobility trace file. The resulting mobility trace covers a period of 30 h, i.e., a whole day plus some margins before and after the evaluated time frame.

To account for diurnal activity patterns, the spawning of pedestrians is thinned out based on the typical hourly vehicle volume of streets[3] [20], which was normalized and depicted in Figure 1. This means that the peak rate of one new pedestrian per second is reached only for hours with peak traffic volume. For the remaining hours, pedestrians can only spawn with a probability corresponding to the relative traffic volume as presented in Figure 1. The resulting mobility trace contains the positions of $43\,447$ unique pedestrians with an average trip length of 1305.24 m. Figure 2 shows the empirical cumulative distribution function (ECDF) of the duration of the simulated trips. Apart from some deviations in the marginal areas, an almost uniform distribution can be observed with an average walking duration of 1379.88 s (ca. 23 min). Thus, according to Little's law, during peak hours, 1379 pedestrians participate in crowdsensing on average, which is roughly 1% of the city's population.
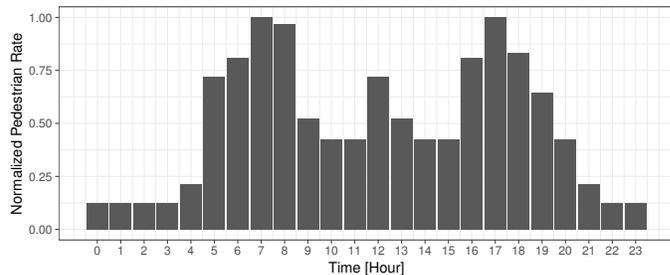

Fig. 1: Normalized pedestrian rate with respect to peak rate. Adapted from the typical hourly vehicle volume of streets [20].
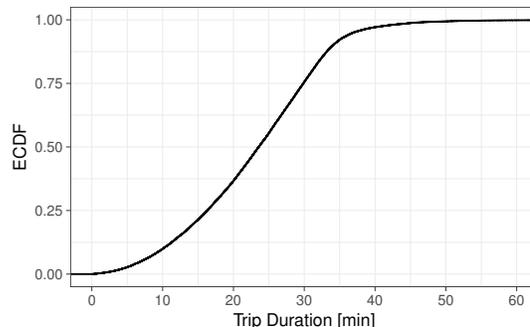

Fig. 2: Distribution of trip duration.

C. Event Detection

Every participating pedestrian is considered a mobile sensor and can detect events. Therefore, the detection range of the mobile sensors has to be modeled. In this study, the map is divided into a regular grid of small cells of width 50 m. The simple assumption is used that an event can be detected if a mobile sensor is in the same cell as the event. Moreover, the detection time of an event can be computed, i.e., the time from the appearance of an event until a mobile sensor covers the event cell. Note that the detection time is 0 if a mobile sensor is co-located in the same cell during the appearance of an event.

IV. RESULTS

The performance of mobile crowdsourcing to detect events can be quantified in terms of detection time and detection probability. Figure 3 shows the ECDF of the detection time. The orange curve represents uncorrelated events with a uniformly distributed location. It can be seen that around 3% of the events are detected immediately with a detection time of 0, which happens when a mobile sensor is already present when an event appears. The ECDF increases fast in the first quartile, but the increase eventually slows down. 57.92% of the events have a detection time larger than 180 min or are never detected. This is not surprising as the map has areas with only a few roads or no roads at all, which significantly lowers the number of potential visitors. Thus, it is difficult or even impossible to detect events that appear in these cells using crowdsensing. In our trace about 53% of all cells we not visited by any mobile sensor at all. The black curve presents the ECDF of the detection time in the case

---

[1]http://www.dlr.de/ts/en/desktopdefault.aspx/tabid-9883/16931_read-41000/
[2]http://sumo.dlr.de/wiki/Vehicle_Type_Parameter_Defaults
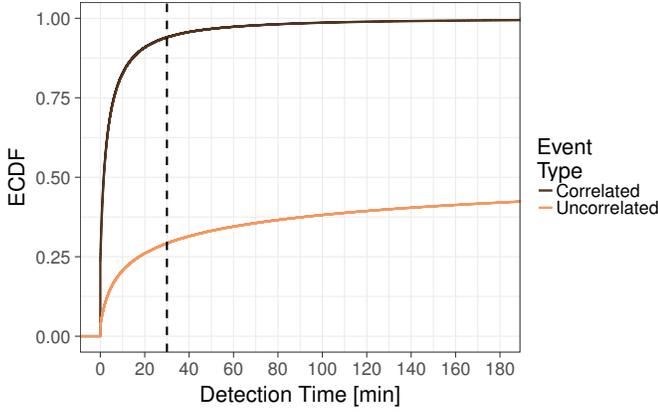[3]https://nacto.org/publication/urban-street-design-guide/design-controls/design-hour/

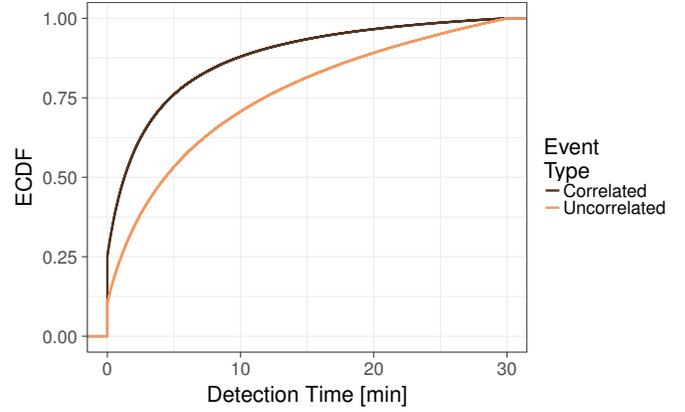Fig. 3: Distribution of detection time for uniform and population based events.

Fig. 5: Distribution of detection time for events with maximum detection time of 30 minutes.
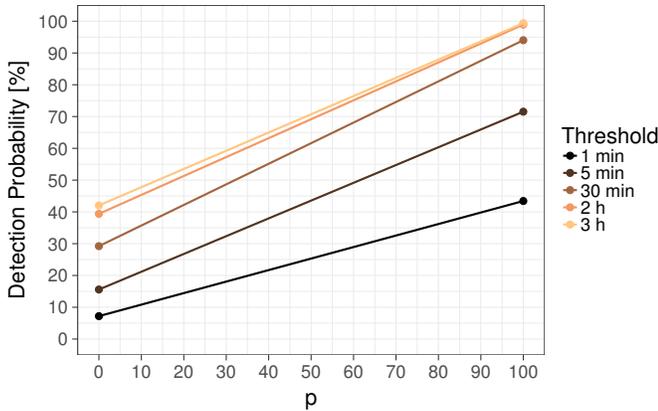

Fig. 4: Detection probabilities for partially correlated events and different maximum detection time thresholds.

of correlated event locations, which are distributed identically to the population density of the mobile trace. The detection times are generally much lower compared to uncorrelated events. 23.8% of the events are detected immediately, and generally, the events are detected much faster. The median is 92 s and only very few events (0.06%) have a detection time larger than 180 min or are never detected.

In real life, events will not be detectable for an infinite amount of time, because they disappear (e.g., a traffic jam dissolves) or become irrelevant (e.g., occasional rain turns into a heavy shower). These events have to be detected within a certain time (e.g., until the next traffic report or weather forecast on a radio channel) to provide useful information. In practice, this maximum detection time might be determined by the type of detected event. However, considering the shapes of the detection time distributions, it can be observed that the detection of uncorrelated events is much more sensitive to setting a maximum detection time than the detection of correlated events.

Figure 4 investigates the trade-offs between the detection probabilities and different maximum detection time thresholds,

which are represented by differently colored lines. The x-axis indicates the percentage p of correlated events, while $(100-p)\%$ of the events are uncorrelated. As the generated partially correlated events are linearly combined from the set of uncorrelated and correlated events and the detections of single events are independent, a linear increase from the detection probability of uncorrelated events dp_{uncorr} in case of $p = 0$ (only uncorrelated events) to the detection probability of correlated events dp_{corr} for $p = 100$ (only correlated events) can be observed for each line. This means that the detection probability dp_p of partially correlated events with percentage p can be computed as

$$dp_p = p\% \cdot dp_{corr} + (100 - p)\% \cdot dp_{uncorr}.$$

The yellow line represents a maximum detection time threshold of 3 h. The detection probabilities of 42.03% for uncorrelated events and 99.39% for correlated events could already be observed at the right margin of Figure 3. When decreasing the maximum detection time threshold down to 1 min (black) to consider very fast event detection only, the detection probabilities fall. However, 7.20% of uncorrelated events and 43.43% of correlated events can still be detected within one minute. This shows that the maximum detection time threshold, which in practice depends on the type of detected event, has a big impact on the performance of mobile crowdsensing.

In the remainder of this work, a maximum detection time of 30 min is assumed. This means, if the detection time is larger than 30 min, an event is considered not detected or missed. In Figure 4, the brown line represents the chosen threshold of 30 min. It can be seen that it makes a good compromise of low detection times and high detection probabilities, which range from 29.22% for $p = 0$ (only uncorrelated events) to 94.02% for $p = 100$ (only correlated events).

Figure 5 shows the resulting distributions when considering only detected events with a maximum detection time of 30 min, i.e., the detection time distribution from Figure 3 truncated at 30 min. It can be seen that still the detection times are shorter for correlated events, which is the expected outcome, but the
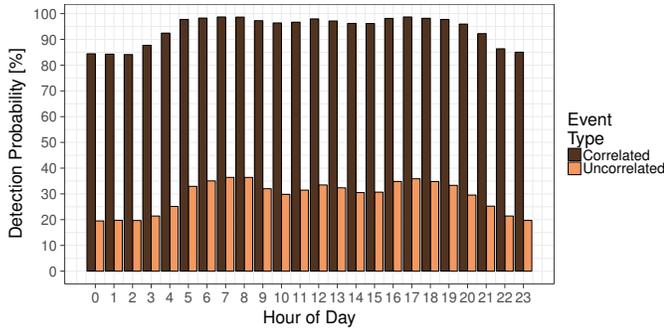
Fig. 6: Hourly detection probability.

(a) Uncorrelated events.

(b) Correlated events.

Fig. 7: Hourly median detection time.

shapes of the ECDFs are more similar. The probabilities of immediate detection are 10.6% for uncorrelated and 25.3% for correlated events, respectively.

As noted above, the allowed maximum detection time of 30 min has a big impact on the detection probability. Figure 6 depicts the hourly detection probability, i.e., the detection probability of an event that appeared during a certain hour of day. Two interesting observations can be made. First, as also indicated by Figure 3, the detection probabilities of correlated events (black) are high and reach a maximum of 98.7%. Even the minimum detection probability of 84.1% during nighttime is considerably high and shows the applicability of crowdsensing for such events. For uncorrelated events (orange), the detection probabilities are much smaller between 19.5% and 36.4%. Thus, for these events mobile crowdsensing has to be complemented by fixed sensors to achieve a serviceable detection probability. The second observation is that the hourly detection probability resembles the crowdsensing participation in Figure 1. This shape is more visible for uncorrelated events as it is superimposed with the generally high detection probability for correlated events. Still, this means that increasing the participation in crowdsensing leads to higher detection probabilities for all kinds of events.

Figure 7 investigates the hourly median detection time, i.e., it shows the median detection time of an event that appeared during that hour of day. The median detection time for uncorrelated events is shown in Figure 7a. It can be seen that the shape resembles the inverse of the crowdsensing participation in Figure 1. The highest median detection times of up to 8.6 min occur during nighttime, as expected, when few people are actively sensing. The median detection times decrease for hours with high crowdsensing participation down to 2.6 min at 7 am and 8 am.

Figure 7b shows the median detection time for correlated events. It has a shape similar to Figure 7a but at a much lower level. At peak hours, the median detection time is 15 s, as many events are immediately detected. The highest median detection time at 1 am is 5.2 min, which is still very low considering the low participation in crowdsensing during that hours. It has to be noted that the detection time distributions have a long tail, which could be seen in Figure 3. Thus, despite low median
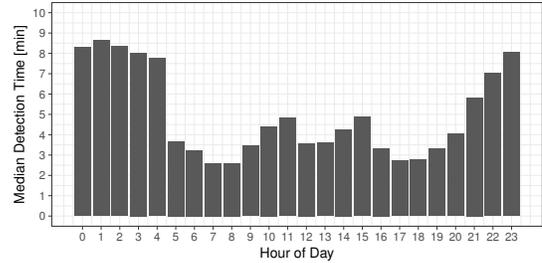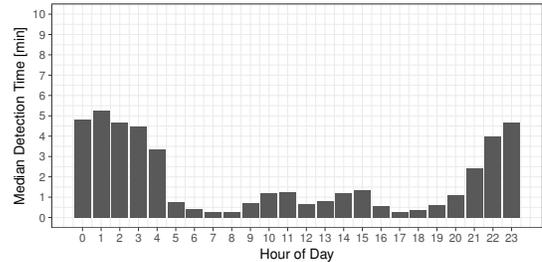
detection times, a considerable amount of events will face much larger detection times or not be detected at all. Nevertheless, most events can be detected very fast by mobile crowdsensing.

V. CONCLUSION

As mobile devices like smartphones or wearables can be easily used as portable sensors, mobile crowdsensing has recently gained increasing attention from research and industry. Especially in case of opportunistic crowdsensing, when participants move and passively sense their environment, possibly unaware of the sensing process, mobile crowdsensing is an economic alternative to installing fixed sensors in smart cities. However, although it might reduce CAPEX and OPEX for fixed sensors, there is a possible reduction of sensor coverage and accuracy due to probabilistic user mobility.

In this work, a simulative performance evaluation of mobile crowdsensing was conducted to investigate these trade-offs for the event detection scenario. A mobility trace of the city of Würzburg with random walks of pedestrians over the course of a single day was used. Thereby, a crowdsensing participation of roughly 1% of the population was simulated. The events appeared uniformly random during 24 h and their locations were distributed over the map of Würzburg, either correlated to the population density, partially correlated, or uncorrelated, i.e., uniformly random. It was shown that the trade-offs between detection time and detection probabilites could be adjusted by different maximum detection time thresholds, although in practice this threshold might be determined by the type of detected event. For the remainder of this work, a maximum detection time of 30 min was assumed.

Uncorrelated events faced long detection times and a rather low detection rate of 29.22%. As the map has areas without

roads, which are never visited by the mobile sensors, events that appear in these cells cannot be detected by mobile crowdsensing. Setting a higher maximum detection time of 180 min can increase the detection rate up to 42.03%, however, the events might disappear or become irrelevant in the meantime. Thereby, the detection of uncorrelated events was much more sensitive to setting a maximum detection time than the detection of correlated events. Correlated events generally showed much shorter detection times and high detection rates of 94.02%. As no events appear in areas without people, high detection rates could be observed even during nighttime. As partially correlated events consisted of $p\%$ correlated and $(100-p)\%$ uncorrelated events, the results for a given percentage p can be interpolated from the marginals.

The results showed that, for correlated events, crowdsensing can achieve almost total coverage of the city. Moreover, it was confirmed that increasing the participation in crowdsensing leads to higher detection rates for all kinds of events. It has to be noted that most detected events could be detected very fast. However, the distribution showed a long tail, which means that a considerable amount of events faced large detection times or could not be detected at all by mobile crowdsensing, especially for uncorrelated events.

In future work, the performance evaluation of the event detection scenario can be improved by considering more realistic mobility traces that also include vehicular mobility. Moreover, the impact of other distributions of event locations can be studied. As the coverage of pure mobile crowdsensing is not perfect yet, hybrid systems of fixed sensors and mobile crowdsourcing will be included and investigated. This adds a monetary dimension as the trade-offs between needs and costs for fixed sensors have to be evaluated. Also users can be paid for sensing sparsely frequented or missing locations/times, which marks the transition from opportunistic to participatory crowdsensing. Finally, new scenarios, such as periodic sensing and continuous sensing, will be tackled.

REFERENCES

[1] D. Krajzewicz, J. Erdmann, M. Behrisch, and L. Bieker, "Recent Development and Applications of SUMO - Simulation of Urban MObility," *International Journal On Advances in Systems and Measurements*, vol. 5, no. 3&4, pp. 128–138, December 2012.

[2] L. Wang, D. Zhang, Y. Wang, C. Chen, X. Han, and A. M'hamed, "Sparse mobile crowdsensing: challenges and opportunities," *IEEE Communications Magazine*, 2016.

[3] N. D. Lane, Y. Chon, L. Zhou, Y. Zhang, F. Li, D. Kim, G. Ding, F. Zhao, and H. Cha, "Piggyback crowdsensing (pcs): energy efficient crowdsourcing of mobile sensor data by exploiting smartphone app opportunities," in *ACM Conference on Embedded Networked Sensor Systems*, 2013.

[4] A. Farshad, M. K. Marina, and F. Garcia, "Urban wifi characterization via mobile crowdsensing," in *IEEE Network Operations and Management Symposium (NOMS)*, 2014.

[5] R. K. Ganti, F. Ye, and H. Lei, "Mobile crowdsensing: current state and future challenges," *IEEE Communications Magazine*, 2011.

[6] B. Guo, Z. Yu, X. Zhou, and D. Zhang, "From participatory sensing to mobile crowd sensing," in *IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM)*, 2014.

[7] N. De Caro, W. Colitti, K. Steenhaut, G. Mangino, and G. Reali, "Comparison of two lightweight protocols for smartphone-based sensing," in *IEEE Symposium on Communications and Vehicular Technology in the Benelux (SCVT)*, 2013.

[8] X. Zhang, Z. Yang, W. Sun, Y. Liu, S. Tang, K. Xing, and X. Mao, "Incentives for mobile crowd sensing: A survey," *IEEE Communications Surveys & Tutorials*, 2016.

[9] I. Koutsopoulos, "Optimal incentive-driven design of participatory sensing systems," in *IEEE Conference on Computer Communications (INFOCOM)*, 2013.

[10] X. Wang, J. Zhang, X. Tian, X. Gan, Y. Guan, and X. Wang, "Crowdsensing-based consensus incident report for road traffic acquisition," *IEEE Transactions on Intelligent Transportation Systems*, 2017.

[11] L. Pournajaf, L. Xiong, V. Sunderam, and S. Goryczka, "Spatial task assignment for crowd sensing with cloaked locations," in *IEEE International Conference on Mobile Data Management (MDM)*, 2014.

[12] S. He, D.-H. Shin, J. Zhang, and J. Chen, "Toward optimal allocation of location dependent tasks in crowdsensing," in *IEEE Conference on Computer Communications (INFOCOM)*, 2014.

[13] H. Roitman, J. Mamou, S. Mehta, A. Satt, and L. Subramaniam, "Harnessing the crowds for smart city sensing," in *International Workshop on Multimodal Crowd Sensing*, 2012.

[14] K. Chakrabarty, S. S. Iyengar, H. Qi, and E. Cho, "Grid coverage for surveillance and target location in distributed sensor networks," *IEEE Transactions on Computers*, 2002.

[15] M. Younis and K. Akkaya, "Strategies and techniques for node placement in wireless sensor networks: A survey," *Ad Hoc Networks*, 2008.

[16] B. Wang, "Coverage problems in sensor networks: A survey," *ACM Computing Surveys (CSUR)*, 2011.

[17] M. Karaliopoulos, O. Telelis, and I. Koutsopoulos, "User recruitment for mobile crowdsensing over opportunistic networks," in *IEEE Conference on Computer Communications (INFOCOM)*, 2015.

[18] D. Zhang, H. Xiong, L. Wang, and G. Chen, "Crowdrecruiter: selecting participants for piggyback crowdsensing under probabilistic coverage constraint," in *ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2014.

[19] H. Ma, D. Zhao, and P. Yuan, "Opportunities in mobile crowd sensing," *IEEE Communications Magazine*, 2014.

[20] National Association of City Transportation Officials, *Urban Street Design Guide*. Island Press, 2013. [Online]. Available: https://nacto.org/publication/urban-street-design-guide/design-controls/design-hour/