This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Short Papers

Sensing Thousands of Visitors Using Radio Frequency

Stijn Denis, *Member, IEEE*, Ben Bellekens, *Member, IEEE*, Maarten Weyn, *Member, IEEE*, and Rafael Berkvens, *Member, IEEE*

***Abstract*—The majority of radio frequency crowd estimation systems are tested with at most a few tens of human individuals. At the same time, they are particularly useful for crowd safety management for events with hundreds or thousands of visitors. We deployed our passive crowd estimation system in two environments at a 400 000-visitor music festival. This paper describes our system's architecture and compares our results to those of an access control system. We found a correlation coefficient of 0.97 between our system and the access control system. Additionally, we show that we can estimate the crowd size correctly up to a few hundred individuals. A system such as ours can provide a direct objective measurement to security personnel who are currently making estimates under stress and influenced by experience, view angle, and occlusions.**

***Index Terms*—Crowd estimation, footfall analytics, passive sensing, radio frequency (RF), sub-GHz.**

I. Introduction

Estimating the number of people present in a particular environment is critical to ensure crowd safety [1]. Many tragic cases exist of fatalities caused by mass panic breaking out in large crowds [2]. Currently, it is often the task of security personnel to estimate if a crowd is too large, but this is a challenging task prone to very human errors. Professional event organizers told us about this problem, but it is also visible in our daily lives: Crowd sizes are often differently reported for political or commercial purposes. Automatically and objectively estimating this number of people is the goal of crowd estimation systems.

The most common crowd estimation system consists of counting the number of people entering and leaving the environment. The environment is required to be accessible only through well-defined entrances and exits. Such crowd estimation systems are commercially available and count using entrance gates, cameras, or reception of wireless signals transmitted by devices carried by the people passing through. These entrance and exit statistics can then be combined to determine the total number of people in the environment. While very useful, these systems operate on a derivative of the actual crowd size and only estimate relatively to when they start operating. If such a system goes offline, it cannot estimate changes in crowd size when resuming operation. We are interested in systems that measure the crowd size directly.

Many direct crowd estimation systems count the number of people visible in a camera image. It has several downsides, however. For large crowd sizes, it is difficult to capture the whole crowd in a single field of view. Furthermore, lighting conditions need to be relatively stable for the accuracy of these systems to remain consistent and the people in the crowd cannot be obscured by smoke or darkness. Finally, the use of vision-based technologies will always have the potential to cause privacy-related issues.

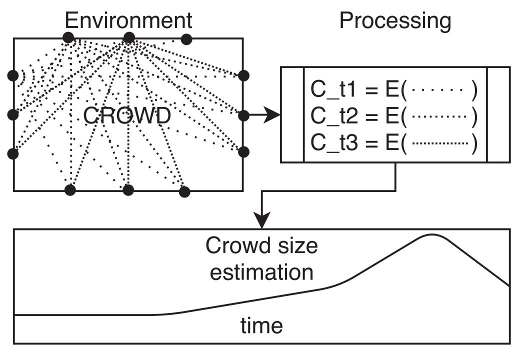


Fig. 1. System architecture: 1) A series of transceivers is placed throughout the environment, broadcasting messages to each other with the direct links passing through the crowd. Three broadcasts are depicted as dotted lines; 2) the signal strength of a broadcast message sent by one transceiver is recorded by all receiving transceivers. This information is passed through the gateway to the processing unit; 3) the crowd size is estimated over time.

We propose to make use of the impact the crowd has on radio frequency (RF) waves in a wireless sensor network (see Fig. 1). These kinds of approaches fall under the umbrella of "passive" or "sensorless" sensing, in which radio transceivers are not only used for communication but also act as sensors. This concept has become a popular topic of research over the last decade and systems utilizing this principle have been developed for device-free localization (DFL) [3], activity recognition [4], fall detection [5], and gesture recognition [6]. Crowd estimation as a research subject can be considered to belong to the "detection" aspect of DFL [7] and quite a few RF-based systems have been developed which focus on this facet [8]–[10]. Many of these approaches solely utilize signal strength measurements, but in recent years, several techniques have appeared which incorporate channel state information (CSI) as well [11]. Such CSI systems use complex hardware, while our proposed system uses low-cost and energy-efficient 868 MHz transceivers that easily cover 100 m in non-line-of-sight conditions, that is to say through a crowd.

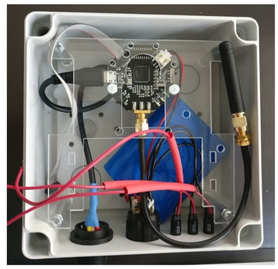
Unfortunately, the vast majority of the systems described in literature are installed in relatively limited environments in which the crowds are no larger than a few tens of people. In this paper, we describe our system architecture during the 2018 edition of the Tomorrowland music festival. As in 2017, we again measured crowd-induced signal attenuation at the "Freedom Stage" and found similar results [12]. The central contribution in this paper is the analysis of measurements in an additional environment, the comfort zone at the "Main Stage," called "Main Comfort." This zone is only accessible after bracelet scanning.

Manuscript received April 28, 2020; revised June 26, 2020; accepted August 10, 2020. *(Corresponding author: Rafael Berkvens.)*

The authors are with IDLab/imec, Faculty of Applied Engineering, University of Antwerp, Antwerpen 2000, Belgium (e-mail: stijn.denis@uantwerpen.be; ben.bellekens@uantwerpen.be; maarten.weyn@uantwerpen.be; rafael.berkvens@uantwerpen.be).

Digital Object Identifier 10.1109/JSYST.2020.3019189

(a)



(b)

Fig. 2. Transceivers, or nodes, used in (a) 'Freedom Stage' and (b) 'Main Comfort' environments. Only one antenna of the 'Freedom Stage' transceiver operates at 868 MHz; the other antenna is not used in this paper.

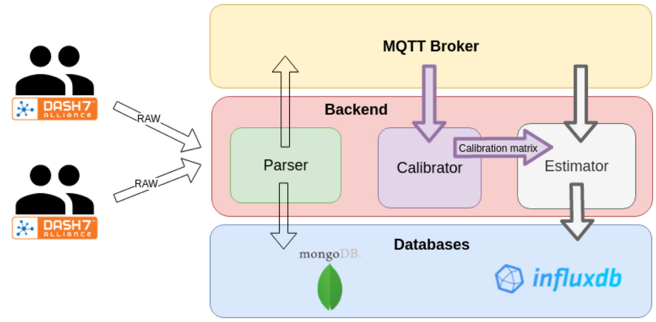


Fig. 3. Data flows within our system's architecture: Purple arrows indicate the calibration flow and grey arrows indicate estimation flows on top of the generic data flow. The schematic shows our technology selection as background information.

We can compare the crowd estimates through this access control system to our system's estimates. The access control system has flaws of its own, however, and we have not been able to quantify the accuracy of these crowd estimates. Despite that, no better count exists for such large crowds.

We structured this short paper as follows. Section II discusses the architecture and deployment of our system. Section III presents our results in the two environments. Section IV discusses these results in a broader context and with regard to what future work remains.

II. System Architecture

The RF-data collection system consists of a series of transceivers, a gateway, and a processing unit. The series of transceivers is installed around and/or within the environment in which the crowd will be present (see Fig. 1). The transceivers are able to broadcast to each other and to the gateway, which passes every measurement toward the processing unit.

Fig. 2 shows the two types of physical transceivers that were deployed in our setups. The first type, Fig. 2(a), consists of an aluminum casing which contained two transceivers with RF front ends that are optimized for, respectively, 433 and 868 MHz communication. We only used the 868 MHz front end in this paper. Because the antennas are on the outside of the node—and therefore clearly visible—we were required by the event organizers to install this type of node in locations which were out-of-sight for the public (e.g., under wooden bars where only staff members could see them). These are the nodes with which we obtained the "Freedom Stage" results.

The second type, Fig. 2(b), has a casing which was entirely closed and made of plastic. In contrast to the previous node type, these antennas are on the inside of the casing, which means that it was less problematic for these nodes to be publicly visible. These nodes contained only an 868 MHz module. These are the nodes with which we obtained the "Main Stage" results.

The general manner in which the transceiver network operates did not significantly change between the experimental measurements described in [12]. Communication occurs in a time slotted manner in which each transceiver in turn broadcasts a data packet toward all other transceivers. The received signal strength (RSS)-value with which a packet is received is saved by each transceiver in an internal list. When the time for a transceiver to broadcast has arrived, the data packet it transmits will contain the values of this list. After the transmission is complete, the internal list is then entirely cleared. Meanwhile, the gateway listens to all communication within the network and passes the received RSS-values on to the processing unit.

The processing unit is responsible for two data flows: The estimation flow and the calibration flow (see Fig. 3). The estimation flow takes the average of the differences of each received signal strength series relative to a calibration matrix of the average RSS-values measured when the environment was (largely) unoccupied. The calibration flow is responsible for creating this square matrix, which has a size equal to the total number of nodes squared. The estimation flow is summarized as

$$a_t = \frac{1}{n^2 - n} \sum_{i=1}^{n^2} (R_m - R_c)_i \tag{1}$$

where a_t is the average attenuation, n is the number of transceivers in the environment, R_m are the received signal strengths from all transceivers to all other transceivers of $n \times n$ —averaged per transceiver over 10 s—and R_c is the calibration matrix of $n \times n$. The fraction is by $n^2 - n$, the number of values without the diagonal values. The diagonal values are all zero after the subtraction and have no meaning since transceivers do not receive their own broadcast. The averages provided by the estimation flow are what we show in this paper to correlate strongly with other crowd estimation measures.

We use linear regression to estimate the crowd size C_t from the average attenuation a_t

$$C_t = a_t x + y \tag{2}$$

where we fit a line through three different sets of training data to determine x and y. One set of training data is selected from a single festival day and all data from the other day is used for evaluation. We combined the resulting errors. If a crowd count estimation was a negative number, it was automatically set to 0.

We repeated our experiment from 2017 at the Freedom Stage [12]. Additionally, we were allowed to install a system at the Main Comfort environment. Fifty-four nodes were installed at about 1.2 m height at the edges of an environment of approximately 25 m by 60 m and allowing 3200 people as per the access control system.

III. Results

Fig. 4 shows the average signal strength attenuation at the "Freedom Stage" environment at the Tomorrowland 2018 music festival, during a single day. Each data point was calculated based on communication

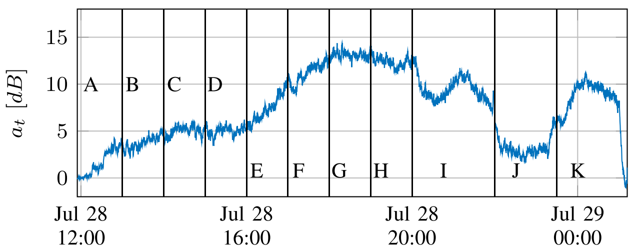


Fig. 4. Average attenuation within our RF-based system in the "Freedom Stage" environment in 2018. The letters separated by vertical lines indicate different artists performing on stage during the day.

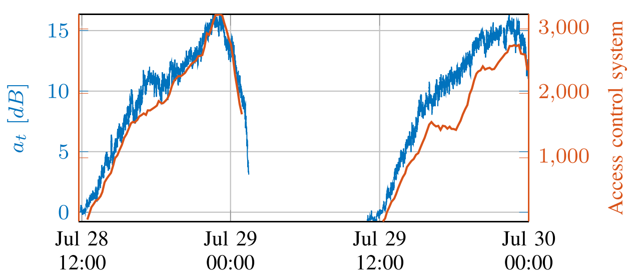


Fig. 5. Average attenuation within our RF-based system and the access control system's crowd size estimate in the "Main Comfort" environment, both scaled between the minimum and maximum value. On July 29, between 14:00 and 16:00, the access control system experienced temporary technical issues.

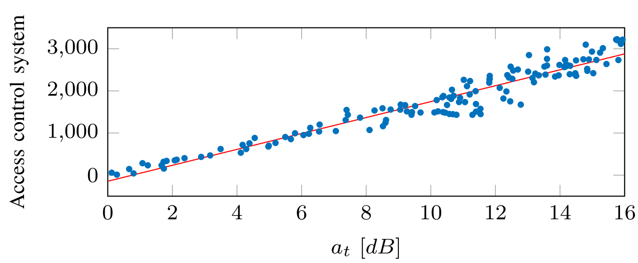


Fig. 6. Correlation between the average attenuation within our RF-based system and the access control system's crowd size estimate in the "Main Comfort" environment. $N = 137$; Pearson's $r = 0.97$.

that occurred within a 10-s interval. This result is comparable to that of our previous work, demonstrating the robustness of the approach [12].

Fig. 5 shows the average signal strength attenuation at the "Main Comfort" environment at Tomorrowland 2018, during two consecutive days. The graphs have been overlaid with the results of the access control system. This system has a number of drawbacks: It registers attendees that go to upper floors of the same building but outside the environment we are measuring; it does not register staff of the festival present in the environment; and it does not register people leaving at the very end of the day. Nevertheless, a comparable trend is visible.

Fig. 6 correlates the average signal strength attenuation at the "Main Comfort" environment with the access control system crowd size estimation. The Pearson's correlation coefficient is 0.97. Note that our calculations only took into account data from the official start of a festival day (12:00) until its official end (01:00 on Saturdays and 00:00 on Sundays). As previously mentioned, people leaving the environment after the end of a festival day were generally not scanned.

Fig. 7 shows the cumulative error distribution for our crowd estimation based on three different linear extrapolation approaches. The

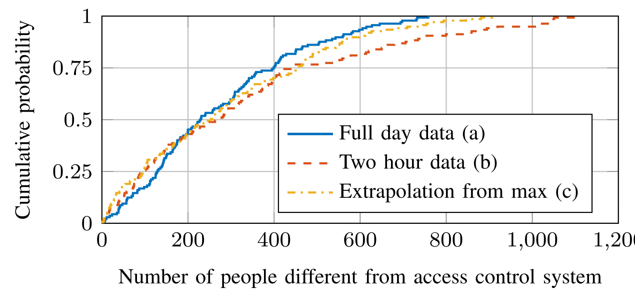


Fig. 7. Cumulative error distributions regarding the difference in estimation with the access control system when using three different training approaches in the "Main Comfort" environment.

first method (a) uses a full day of training data to estimate x and y of (2) and validates on the other day. The median is 228.9650, the mean is 270.2352, the standard deviation is 184.3157, and 95th percentile is 622.9059. The second method (b) uses two hour of training data of one day and validates on the other day. The median is 264.0075, the mean is 328.9839, the standard deviation is 285.3301, and 95th percentile is 979.8963. The third method (c) uses only two data points: Zero at the start of day, and the max values of one day, validating on the other day. The median is 253.4367, the mean is 284.9903, the standard deviation is 224.4483, and 95th percentile is 733.1686. The error is calculated assuming that the access control system is the correct value, despite its drawbacks.

IV. Discussion and Conclusion

We explained how we determine the average attenuation of the RF-network in an environment. The experiments in the "Freedom Stage," both in 2017 (see our previous work [12]) and in 2018 (see Fig. 4), indicate that this value provides information regarding the size of the crowd that is currently present in the environment: A direct crowd size estimate as outlined in Section I. The contribution in this paper consists of translating the average attenuation into an actual crowd size estimate.

Fig. 5 compares the graphs of the average attenuation and the number of people present according to the scanning system during the final two days of the festival. Strong correlations can be observed when we scale both graphs between their respective minimum and maximum values. On the last festival day, between 14:00 and 16:00, the two graphs diverge, with the access control system reporting a steady crowd size and our system reporting an increasing average attenuation. We suspect that the divergence is due to people moving to the free food served on floors outside our measurement area, but we did not remove this data discrepancy in what follows.

As outlined in the previous section, we use three different training methodologies to obtain x and y in (2). Interestingly, even a very basic training method that uses just the empty and full environment as the only two data points obtains very similar results to the more involved training methods. The full day training method, however, seems slightly less prone to outliers. The errors are in the order of 2–300 people, with a deviation of another 200, with a 95th percentile error reaching 900; it would seem that such accuracy is unsatisfactory because that 95th percentile is about a third of the maximum number estimated by the access control system. Of course, we have stressed that our validation for this setup was not perfect. The scanning system has a number of flaws as validation: Counting all people rather than just those on the floor we measured; not counting staff; and not counting people leaving at the end of the day. The scanning system provides a crowd size every 15 min, while our system currently updates every 10 s. Moreover, professional event organizers are used to deal with much larger deviations and still have to try to keep the crowd safe. The organizers already indicated interest in our system, and we see potential for further improvement.

The attenuation values are determined by calculating the differences between current RSS measurements and RSS measurements that were performed when the environment was (largely) unoccupied by human individuals. This means that it is necessary to delineate a calibration period in time during which the environment is both static and unoccupied in order to perform calibration measurements. Furthermore, during medium- to long-term setups in which the position of generally static objects in the environment can change over the course of the measurements (e.g., slightly changed positions of the tables in the Main Comfort environment from one day to the next), these measurements need to be regularly updated. Early analyses indicate that it is possible to automatically determine the time intervals which are most suitable for these measurements to occur, without needing constant visual verification that the environment is truly empty. We observed that by analyzing the mean variance experienced by communication links, periods of time in which the environment was largely static—and which were therefore prime candidates to be used as calibration periods—could be detected. This observation forms the basis for an autocalibration system which we are currently developing. Re-calibrating slope parameter x is more challenging because it depends on the network layout, the environment, and weather conditions such as humidity and precipitation.

Currently, we do not have numerical validation data for any other experimental setup; the fact that it is difficult to obtain underlines the need for an automatic objective crowd size estimation system. Our system's crowd size estimates, however, do typically lead to a realistic view of changing crowd sizes in other setups. This is so far only confirmed by manual analysis of camera images (when available) and the expert opinion of event organizers.

There are a number of future research questions. For example, we suspect that the estimation error decreases when the node density increases, but we will need to validate this in smaller environments with hundreds of people. Moreover, we need to quantify the impact of people moving close to the sensing nodes and the impact of changing environment topology or furniture after calibration.

To summarize, our system was installed in a real-life setup at a large scale, commercial festival environments containing up to thousands of individuals, in contrast to similar passive RF-based sensing approaches. In the environment for which we had access to validation data, the correlation coefficients between the variables representing this average attenuation and the actual number of people present is 0.97. A linear regression based approach for crowd estimation led to median estimation errors between 184 and 227 individuals, within an environment with approximate maximum occupancy of 3200 people. The real-time RF attenuation based data we were able to provide to the event organizers was already considered to be particularly useful in a security and safety context at large-scale events. We hope that direct crowd size estimation system like ours can provide an objective measure for commercial, political, and safety decisions.

References

[1] M. W. Aziz, F. Naeem, M. H. Alizai, and K. B. Khan, "Automated solutions for crowd size estimation," *Social Sci. Comput. Rev.*, vol. 36, no. 5, pp. 610–631, 2018.

[2] S. A. M. Saleh, S. A. Suandi, and H. Ibrahim, "Recent survey on crowd density estimation and counting for visual surveillance," *Eng. Appl. Artif. Intell.*, vol. 41, pp. 103–114, 2015.

[3] O. Kaltiokallio, R. Jäntti, and N. Patwari, "ARTI: An adaptive radio tomographic imaging system," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 7302–7316, Aug. 2017.

[4] B. Wei, W. Hu, M. Yang, and C. T. Chou, "Radio-based device-free activity recognition with radio frequency interference," in *Proc. 14th Int. Conf. Inf. Process. Sensor Netw.*, ACM, 2015, pp. 154–165.

[5] S. Kianoush, S. Savazzi, F. Vicentini, V. Rampa, and M. Giussani, "Device-free RF human body fall detection and localization in industrial workplaces," *IEEE Internet Things J.*, vol. 4, no. 2, pp. 351–362, Apr. 2016.

[6] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-home gesture recognition using wireless signals," in *Proc. 19th Annu. Int. Conf. Mobile Comput. Netw.*, ACM, 2013, pp. 27–38.

[7] M. Youssef, M. Mah, and A. Agrawala, "Challenges: Device-free passive localization for wireless environments," in *Proc. 13th Annu. Int. Conf. Mobile Comput. Netw.*, ACM, 2007, pp. 222–229.

[8] Y. Yuan, J. Zhao, C. Qiu, and W. Xi, "Estimating crowd density in an RF-based dynamic environment," *IEEE Sensors J.*, vol. 13, no. 10, pp. 3837–3845, Oct. 2013.

[9] S. Depatla, A. Muralidharan, and Y. Mostofi, "Occupancy estimation using only WiFi power measurements," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 7, pp. 1381–1393, Jul. 2015.

[10] S. Di Domenico, M. De Sanctis, E. Cianca, P. Colucci, and G. Bianchi, "LTE-based passive device-free crowd density estimation," in *Proc. 2017 IEEE Int. Conf. Commun.*, IEEE, 2017, pp. 1–6.

[11] S. Di Domenico, M. De Sanctis, E. Cianca, and G. Bianchi, "A trained-once crowd counting method using differential WiFi channel state information," in *Proc. 3rd Int. Workshop Phys. Anal.*, ACM, 2016, pp. 37–42.

[12] S. Denis, R. Berkvens, B. Bellekens, and M. Weyn, "Large scale crowd density estimation using a sub-GHz wireless sensor network," in *Proc. 2018 IEEE 29th Annu. Int. Symp. Pers. Indoor Mobile Radio Commun.*, IEEE, 2018, pp. 849–855.